UNIVERSITY OF ALBERTA

THE CONSTRUCTION AND REVISION OF TEST ITEMS REFERENCED TO SELECTED FIFTH-GRADE MUSIC READING ABILITIES

by

SHEILA JOAN SCOTT ©

A Dissertation submitted to the Faculty of Graduate Studies and Research in partial
fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELEMENTARY EDUCATION

EDMONTON, ALBERTA

SPRING, 1997

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced with the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-21630-6

Canadä

University of Alberta

Library Release Form

Name of Author: Sheila Joan Scott

Title of Dissertation: The Construction and Revision of Test Items Referenced to Selected Fifth-Grade Music Reading Abilities

Degree: Doctor of Philosophy

Year This Degree Granted: 1997

Permission is hereby granted to the University of Alberta to reproduce single copies of this dissertation and to lend or sell such copies for private, scholarly, or scientific research purposes.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Sheila J. Scott

1446 Taylor Drive
Swift Current, Saskatchewan
S9H 1M9

DATED December 11, 1996

UNIVERSITY OF ALBERTA

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled THE CONSTRUCTION AND REVISION OF TEST ITEMS REFERENCED TO SELECTED FIFTH-GRADE MUSIC READING ABILITIES submitted by SHEILA JOAN SCOTT in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

R J de Frece
Dr. Robert de Frece (Supervisor)

Dr. G. J. Fishburne
Dr. Graham Fishburne

Amanda Montgomery
Dr. Amanda Montgomery

Rogers
Dr. W. Todd Rogers

Robert Ware
Dr. Robert Ware

Robert Ware for
Dr. Randall Moore (External Examiner)
University of Oregon

Date: Sept. 24, 1996

DEDICATION

This dissertation is dedicated to my parents Ewart and Joyce and my sister Heather who have always supported by academic endeavors.

ABSTRACT

The purposes of this study were (1) to identify the various ways in which the ability to read and/or write tonal melodies in traditional music notation is addressed in selected teaching materials and (2) to devise pencil-and-paper test items by which fifth-grade students may demonstrate performance in terms of these items. The first problem was addressed by analyzing selected educational resources and identifying characteristics of melody related to music literacy. The second problem was addressed in three stages: 1. the development of initial test items based on the professional judgement of the test developer; 2. the assessment of the relevance and representativeness of these items based on the professional judgement of subject-area specialists; and, 3. the examination of student responses based on data obtained in six field studies, each which employed a revised version of the test.

Eleven characteristics of melody identified in the analysis of educational resources were transformed into 13 instructional objectives. Each objective was represented by two test items, resulting in an initial test version of 26 items. Two issues were addressed in test administration: first, the function of the test items as revealed in classical item analysis (Field Studies 1, 3, 5, and 6) and, second, the strategies employed by students when solving these items as demonstrated through the analysis of student protocols (Field Studies 2 and 4). The majority of item revisions were assigned to five classifications—melody, rhythm, target note, directions to students, and practise questions.

Test Version 6 (22 items), administered to 214 grade 5 students in the Spring of 1995, was arranged in two main sections: first, 14 items were presented both aurally (soprano voice) and visually (music notation); and, second, 8 items were presented in visual mode only. An oblique factor analytic solution approximating simple structure yielded three factors: 1. aural/visual; 2. visual; and 3. contour. Aggregated scores ranged from 1 to 21. The mean was 11.25 and the standard deviation was 5.0 (4.4%).

The internal consistency of .84 (coefficient alpha) exceeded the value of .80 generally accepted for this type of measure.

ACKNOWLEDGMENTS

First and foremost I am thankful to Allah the Almighty (God) for his Grace and Mercy.

I would like to express my sincere thanks to my supervisor, Dr. N. Rajaratnam for his guidance, encouragement, valuable advice and support throughout the course of this study. His strong interest and enthusiasm in the subject and encouraging support was very inspiring and helped to make this research possible.

Thanks are due to Mr. Sheldon Lovell for setting up the experimental arrangements at different stages of this study.

I would like to extend my thanks to Mr. Chris Katopodis, Department of Fisheries and Oceans, Government of Canada, for his valuable support and suggestions.

I wish to thank all my colleagues and professors in the Water Resources group for their friendship and useful discussions on the subject, especially Dr. S. Wu for helping to set-up the data acquisition system and Dr. F. Ahmed for his initial input.

I am very grateful to my wife for her support, encouragement and understanding. Thanks to my family I was able to finish this research program. Special thanks to my parents for their support and encouragement.

I would like to acknowledge the financial support of the Ministry of Culture and Higher Education of the Islamic Republic of Iran in the form of a scholarship and a research assistantship through an operating grant to my supervisor from the Department of Fisheries and Oceans, Government of Canada.

Acknowledgements

I gratefully acknowledge those persons who contributed to this study. Appreciation is extended to my advisor, Dr. Robert de Frece, who assisted in the numerous stages of research essential to the completion of this extended project. I am indebted to Dr. W. Todd Rogers whose expertise in research design and measurement and evaluation was integral to the development of the six field studies on which this study was founded. I also wish to thank Dr. Robert Ware for supporting my work throughout the course of this project. Additionally, appreciation is extended to members of my examining committee, Dr. Amanda Montgomery and Dr. Graham Fishburne for participating in the final defense of this document and to Dr. Randall Moore for agreeing to be the external examiner.

This research would not have been possible without the assistance of teachers in the field. In this respect, I wish to thank the music teachers who participated in this study by completing questionnaires and by allowing the test items developed in this study to be administered to their students. In addition, I wish to thank Carolyn Wentzel and Joyce Ndalichako, consultants in the Centre for Research in Applied Measurement and Evaluation, who assisted with the generation of computer printouts which formed part of the basis for data analysis.

I am grateful for the opportunity to know and work with graduate students at the University of Alberta. In particular, I would like to thank Karyn Cooper and Marilyn Kerley for their intellectual insight and friendship during the completion of this document and beyond. And finally, a sincere thank you is extended to Suzannne Derbecker, Jennifer Lock, Irene Mah, Landon Modien, and Morag Northy for their unwavering confidence in my pursuit for academic excellence.

TABLE OF CONTENTS

Chapter Page

1. THE NEED FOR THE STUDY 1
    Background 1
    Problem and Sub-Problems 3
    Delimitations 4
    Definition of Terms 4
        Extended Definitions 5
            Capacity 5
            Ability 6
            Aptitude 7
            Achievement 9
        General Definitions 9
            Tonality 9
            Melody 10
            Temporal Organization 10
            Audiation 10
            Music Literacy 10
            Notational Discrimination 11
    Organization of the Thesis 11

2. RELATED LITERATURE 12
    Overview 12
    Aural Perception Within a Tonal Context 12
    Tonal and Durational Relationships 18
    A Developmental Perspective 21
    Music Reading 26
        Psychological Studies 26
        Educational Studies 29
    Standardized Music Tests Incorporating Tasks of Music Reading 31
        Test of Musicality 32
        Farnum Test of Music Notation 33
        Music Achievement Tests 34

Chapter Page

Iowa Tests of Music Literacy 35
Summary 37
Selected Published Music Tests and Gender Bias 37
Claims of Test Authors 37
Additional Findings 39
Summary 40

3 DEVELOPMENT OF THE INITIAL TEST ITEMS 41
Overview 41
Initial Test Development 42
Item Construction 42
Educational Representation of Music Reading 42
Melodic Characteristics 42
Item Specifications 42
Number of Items 45
Melodies Composed for Item Development 45
Educational Sources 45
Scale Structure 45
Metric Structure 47
Rhythmic Structure 47
Melodies Composed for this Study 49
Scale Structure 49
Metric Structure 49
Rhythmic Structure 49
Psychological Representation 49
Song Text 52
Item Format 52
Sequential Organization of Test Items 54
Aural Presentation 55
Mode of Response 55
Vocal Component 55
Time Allocated for Item Completion 56
Content-Related Evidence of Validity 58
Questionnaire Responses From Elementary Music Specialists 58

Chapter Page

Iowa Tests of Music Literacy 35
Summary 37
Selected Published Music Tests and Gender Bias 37
Claims of Test Authors 37
Additional Findings 39
Summary 40

3 DEVELOPMENT OF THE INITIAL TEST ITEMS 41
Overview 41
Initial Test Development 42
Item Construction 42
Educational Representation of Music Reading 42
Melodic Characteristics 42
Item Specifications 42
Number of Items 45
Melodies Composed for Item Development 45
Educational Sources 45
Scale Structure 45
Metric Structure 47
Rhythmic Structure 47
Melodies Composed for this Study 49
Scale Structure 49
Metric Structure 49
Rhythmic Structure 49
Psychological Representation 49
Song Text 52
Item Format 52
Sequential Organization of Test Items 54
Aural Presentation 55
Mode of Response 55
Vocal Component 55
Time Allocated for Item Completion 56
Content-Related Evidence of Validity 58
Questionnaire Responses From Elementary Music Specialists 58

Chapter Page

Description of the Sample 58
Structure of the Questionnaire 59
Summary of Responses 61
  Reading Level 61
  Item Clarity 61
  Item Relevance 61
  Item Representativeness 61
Cluster-By-Cluster Analysis 66
  Cluster 1 (items 1 and 2) 66
    Questionnaire Responses 66
    Suggestions For Item Revision 66
  Cluster 2 (items 3 and 4) 66
    Questionnaire Responses 66
    Suggestions For Item Revision 67
  Cluster 3 (items 5 and 6) 68
    Questionnaire Responses 68
    Suggestions For Item Revision 68
  Cluster 4 (items 7 and 8) 69
    Questionnaire Responses 69
    Suggestions For Item Revision 69
  Cluster 5 (items 9 and 10) 70
    Questionnaire Responses 70
    Suggestions For Item Revision 70
  Cluster 6 (items 11 and 12) 70
    Questionnaire Responses 70
    Suggestions For Item Revision 71
  Cluster 7 (items 13 and 14) 72
    Questionnaire Responses 72
    Suggestions For Item Revision 72
  Cluster 8 (items 15 and 16) 72
    Questionnaire Responses 72
    Suggestions For Item Revision 72
  Cluster 9 (items 17 and 18) 73
    Questionnaire Responses 73

Chapter Page

Suggestions For Item Revision 73
Cluster 10 (items 19 and 20) 73
Questionnaire Responses 73
Suggestions For Item Revision 74
Cluster 11 (items 21 and 22) 74
Questionnaire Responses 74
Suggestions For Item Revision 74
Cluster 12 (items 23 and 24) 74
Questionnaire Responses 74
Suggestions For Item Revision 74
Cluster 13 (items 25 and 26) 75
Questionnaire Responses 75
Suggestions For Item Revision 75
Time Allocated for Task Completion 76
Trial Administration 76
Data Analysis 76
Chapter Summary 78

4 FIELD STUDY 1: TEST VERSION 1 79
Overview 79
Description of the Sample 79
Overview of Studies 1 Through 6 79
Field Study 1 80
Item Analysis 81
Method 81
Item Difficulty 81
Item Discrimination 82
Field Study 1 84
Item-By-Item Analysis 84
Items 1 and 2 84
Indices of Difficulty and Discrimination 84
Item Revision 84
Items 3 and 4 87
Indices of Difficulty and Discrimination 88

Chapter Page

Item Revision 89
Items 5 and 6 90
Indices of Difficulty and Discrimination 90
Item Revision 92
Items 7 and 8 92
Indices of Difficulty and Discrimination 92
Item Revision 94
Items 9 and 10 94
Indices of Difficulty and Discrimination 94
Item Revision 94
Items 11 and 12 95
Indices of Difficulty and Discrimination 96
Item Revision 96
Items 13 and 14 96
Indices of Difficulty and Discrimination 97
Item Revision 97
Items 15 and 16 98
Indices of Difficulty and Discrimination 98
Item Revision 98
Items 17 and 18 98
Indices of Difficulty and Discrimination 99
Item Revision 100
Items 19 and 20 100
Indices of Difficulty and Discrimination 101
Item Revision 102
Items 21 and 22 102
Indices of Difficulty and Discrimination 102
Item Revision 104
Items 23 and 24 104
Indices of Difficulty and Discrimination 104
Item Revision 105
Items 25 and 26 105
Indices of Difficulty and Discrimination 105
Item Revision 105

Chapter Page

Chapter Summary 106

5 FIELD STUDY 2: TEST VERSION 2 107
  Introduction 107
  Method 107
    Instrument 107
    Student Sample 107
    Interviews 108
  Analysis of Probes 109
    Categorizations 109
      Open-ended Reminders 110
      Requests for Additional Information 110
      Verbal Reminder 111
      Attention to Options 112
      Implications for Item Revision 113
  Data Analysis 113
    Item-By-Item Analysis 115
      Item 1 115
        Correct Response 115
        Incorrect Response 116
        Item Revision 116
      Item 4 117
        Correct Response 117
        Incorrect Response 117
        Item Revision 117
      Item 5 118
        Correct Response 118
        Incorrect Response 119
        Item Revision 119
      Item 7 119
        Correct Response 120
        Incorrect Response 120
        Item Revision 120
      Item 9 120

Chapter Page

    Correct Response 120
    Incorrect Response 120
    Item Revision 121
  Item 11 121
    Correct Response 121
    Incorrect Response 122
    Item Revision 122
  Item 13 122
    Correct Response 123
    Incorrect Response 123
    Item Revision 123
  Item 15 123
    Correct Response 123
    Incorrect Response 124
    Item Revision 124
  Item 17 124
    Correct Response 125
    Incorrect Response 125
    Item Revision 125
  Item 19 125
    Correct Response 125
    Incorrect Response 126
    Item Revision 126
  Item 21 126
    Correct Response 126
    Incorrect Response 127
    Item Revision 128
  Item 24 129
    Correct Response 129
    Incorrect Response 129
    Item Revision 129
  Item 26 130
    Correct Response 130
    Incorrect Response 130

Chapter Page

Item Revision 131
Chapter Summary 131

6 FIELD STUDY 3: TEST VERSION 3 133
Overview 133
Item Revisions 134
Items 1 and 2 137
Indices of Difficulty and Discrimination 137
Item Revision 137
Item 4 138
Indices of Difficulty and Discrimination 138
Item Revision 139
Items 5 and 6 140
Indices of Difficulty and Discrimination 140
Item Revision 140
Items 13 and 14 142
Indices of Difficulty and Discrimination 142
Item Revision 142
Items 15 and 16 143
Indices of Difficulty and Discrimination 143
Item Revision 143
Items 17 and 18 144
Indices of Difficulty and Discrimination 145
Item Revision 145
Items 19 and 20 146
Indices of Difficulty and Discrimination 146
Item Revision 147
Items 23 and 24 150
Indices of Difficulty and Discrimination 150
Item Revision 151
Summary Description 152
Summary Test Statistics 152
Chapter Summary 152

Chapter Page

7. MODE OF PRESENTATION (FIELD STUDIES 4 AND 5) 153
   - Overview 153
   - Field Study 4: Protocol Analysis 153
      - Introduction 153
      - Results 155
   - Field Study 5: Group Administration 156
      - Overview 156
      - Mode of Presentation 156
      - Assessment of Visual Mode Only 156
      - Item Revision 158
         - Items 1 and 2 158
            - Indices of Difficulty and Discrimination 161
            - Item Revision 161
         - Item 4 162
            - Indices of Difficulty and Discrimination 162
            - Item Revision 162
         - Items 5 and 6 163
            - Indices of Difficulty and Discrimination 163
            - Item Revision 164
         - Items 9 and 10 164
            - Indices of Difficulty and Discrimination 164
            - Item Revision 164
         - Items 15 and 16 165
            - Indices of Difficulty and Discrimination 165
            - Item Revision 166
         - Items 17 and 18 166
            - Indices of Difficulty and Discrimination 166
            - Item Revision 166
         - Items 23 and 24 167
            - Indices of Difficulty and Discrimination 167
            - Item Revision 167
   - Chapter Summary 168

8 FIELD STUDY 6: TEST VERSION 6 170
- Overview 170
- Description of the Sample 170
- Results 171
  - Item Revision 171
    - Item 11 171
      - Indices of Difficulty and Discrimination 171
      - Item Revision 171
    - Item 13 174
      - Indices of Difficulty and Discrimination 174
      - Item Revision 175
    - Item 16 175
      - Indices of Difficulty and Discrimination 175
      - Item Revision 175
    - Items 21 and 22 175
      - Indices of Difficulty and Discrimination 175
      - Item Revision 176
  - Summary Test Statistics 176
- Test Validity 177
  - Students' Opportunity to Learn 177
    - Structure of the Questionnaire 177
    - Summary of Responses 178
      - Characteristic of Melody 178
      - Test Items 181
      - Time of Instruction 182
      - Instructional Materials 182
    - Summary 182
  - Influence of Formal Music Instruction and Gender 183
  - Factor Structure 185
    - Overview 185
    - Method 186
    - Description of the Factors 186
      - Factor 1: Aural/Visual 188
      - Factor 2: Visual 188
      - Factor 3: Contour 189

Chapter Page

Complex Variables 189
Item 11 189
Item 13 190
Item 14 190
Item 20 191
Nonloading Item 191
Factor Correlation 192
Summary 192
Chapter Summary 193

9 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS 195
Summary of Findings 195
Limitations 197
Implications for Future Research 197
Implications for Practise 198

REFERENCES 200

APPENDICES

A. INITIAL TEST ITEMS 212
B. QUESTIONNAIRE COMPLETED BY SUBJECT AREA SPECIALISTS: ILLUSTRATION OF STRUCTURE 223
C. LETTERS TO PARTICIPANTS 228
D. TEST VERSION 1 231
E. TEST VERSION 2 242
F. TEST VERSION 3 248
G. FIELD STUDY 3: ITEM-BY-ITEM-ANALYSIS FOR ITEMS NOT REVISED 258
H. TEST VERSION 4S 266
I. FIELD STUDY 4: ITEM-BY-ITEM ANALYSIS 270
J. TEST VERSION 5 281
K FIELD STUDY 5: ITEM-BY-ITEM ANALYSIS FOR ITEMS NOT REVISED 292
L. TEST VERSION 6 297

Appendix Page

M. FIELD STUDY 6: ITEM-BY-ITEM ANALYSIS FOR ITEMS NOT REVISED 307
N. QUESTIONNAIRE COMPLETED BY MUSIC TEACHERS: ILLUSTRATION OF STRUCTURE 313
O. FORMAL MUSIC EXPERIENCE 318
P. FACTOR STRUCTURE: PATTERN MATRIX 323
Q. ETHICS APPROVAL 325

LIST OF TABLES

Table Page

1. Melodic Characteristics Introduced by the End of Grade 4 43
2. Melodic Characteristics Related to Sources Consulted 44
3. Melodic Characteristics Transformed to Instructional Objectives 46
4. Grades at which Scales are First Introduced in the Educational Literature .. 47
5. Grades at which Meters are First Introduced in the Educational Literature . 48
6. Grades at which Rhythms are First Introduced in the Educational Literature 48
7. Scalar, Metric, and Rhythmic Representation of Test Items 50
8. Categorization of Test Items: Item format and mode of response 54
9. Period of Silence Allocated for Item Completion 57
10. Questionnaire Judges: Educational Background, Years of Experience and Grades Taught 60
11. Questionnaire Responses: Reading Level (question 1) 62
12. Questionnaire Responses: Item Clarity (question 2) 63
13. Questionnaire Responses: Item Relevance (question 3) 64
14. Questionnaire Responses: Item Representativeness (question 4) 65
15. Time Allocated and Times Heard: Trial Administration and Test Version 1 77
16. Summary of Field Studies 80
17. Field Study 1: Indices of Difficulty and Discrimination 85
18. Field Study 3: Indices of Difficulty and Discrimination 135
19. Field Studies 5 - Visual Only and 3 - Visual-Aural: Indices of Difficulty and Discrimination for Items 17 to 24 157
20. Field Study 5: Indices of Difficulty and Discrimination 159
21. Field Study 6: Indices of Difficulty and Discrimination 172
22. Questionnaire Responses: Characteristic of Melody (question 1) 179
23. Questionnaire Responses: Test Items (question 2) 179
24. Questionnaire Responses: Time Period (question 3) 180
25. Questionnaire Responses: Instructional Materials 180
26. Cell Distribution 184
27. Analysis of Variance: Total Score (Dependent Variable) 184
28. Factor Analysis: Oblique Solution (Pattern Matrix) 187
29. Factor Correlation Matrix 188

LIST OF FIGURES

Figure Page

1. Item 17: Multiple-choice format 53
2. Item 5: Written response format 53
3. Aural cue 56
4. W-shaped melody 67
5. Circled notes as illustrated in directions to students 68
6. Item 6: Visual response versus aural response 69
7. Item 9: Comparison of Initial Test Version and Test Version 1 71
8. Item 18: Comparison of Initial Test Version and Test Version 1 73
9. Item 25: Comparison of Initial Test Version and Test Version 1 75
10. Formula for item difficulty 82
11. Formula for the point-biserial index of correlation 83
12. Item 1: Diagrammatic and notational representations of contour 87
13. Repeated notes: Student interpretation 88
14. Repeated notes: Expected theoretical interpretation 88
15. Item 4: Student response 89
16. Item 3: Revised form 89
17. Item 5: Comparison of melodies employed in Test Versions 1 and 2 91
18. Item 6: Melodic material 91
19. Item 7: Comparison of melodies employed in Test Versions 1 and 2 93
20. Item 8: Melodic material 93
21. Item 9: Student response 95
22. Item 11: Target note 95
23. Item 12: Target note 95
24. Item 13: Target notes 96
25. Item 14: Comparison of Test Versions 1 and 2 97
26. Item 15: Comparison of Test Versions 1 and 2 99
27. Item 18: Step-wise and disjunct motion 100
28. Item 17: Comparison of Test Versions 1 and 2 100
29. Item 19: Melodic Material 101
30. Item 20: Comparison of Test Versions 1 and 2 102
31. Item 21: Structural flaw revealed in option C 103
32. Item 25: Comparison of Test Versions 1 and 2 105

Figure Page

33. Item 26: Comparison of Test Versions 1 and 2 106

34. Item 1: Comparison of Test Versions 2 and 3 116
35. Item 4: Comparison of Test Versions 2 and 3 118
36. Item 5: Comparison of Test Versions 2 and 3 120
37. Item 19: Comparison of Test Versions 2 and 3 127
38. Item 21: Directions to students 128
39. Item 21: Comparison of Test Versions 2 and 3 for option C 128
40. Item 24: Comparison of Test Versions 2 and 3 130
41. Item 26: Melody employed in dictation task 130
42. Item 26: Comparison of Test Versions 2 and 3 132
43. Item 1: Comparison of Test Versions 3 and 4 138
44. Item 2: Comparison of Test Versions 3 and 4 138
45. Item 4: Comparison of Test Versions 3 and 4 139
46. Musicanada: Steps and skips 141
47. Item 5: Comparison of Test Versions 3 and 4 141
48. Item 6: Comparison of Test Versions 3 and 4 142
49. Item 16: Aural and visual representation 144
50. Item 15: Aural and visual representation 144
51. Item 17: Linear and complex contour 146
52. Item 18: Linear and complex contour 146
53. Item 19: Comparison of option C in Test Versions 3 and 4 148
54. Item 19: Comparison of option B in Test Versions 3 and 4 148
55. Item 20: Comparison of option A in Test Versions 3 and 4 149
56. Item 19: Test Version 4 149
57. Item 20: Test Version 4 150
58. Item 23: Comparison to instructions to students in Test Versions 3 and 4 151
59. Home tone: Melodies presented in visual and aural/visual mode 155
60. Item 1: Comparison of target notes in Test Versions 5 and 6 161
61. Item 2: Comparison of target notes in Test Versions 5 and 6 162
62. Item 4: Target notes 163
63. Item 9: Comparison of aural representations in Test Versions 5 and 6 164
64. Item 10: Comparison of Test Versions 5 and 6 165
65. Item 23: Comparison of directions to students in Test Versions 5 and 6 .. 167

Figure Page

66. Item 11: Option C 174

67. Item 11: Test Version 7 174

68. Item 16: Melody employed in Test Versions 6 and 7 176

69. Items 11, 9, and 12: Correct response 190

70. Item 14: Correct response 191

CHAPTER 1

THE NEED FOR THE STUDY

Background

Psychologists and musicians have pondered the fundamental nature of music understanding since the turn of the century. This has precipitated the development of numerous standardized tests of music aptitude (Drake, 1957; Gaston, 1957; Gordon, 1965, 1979b, 1982; Kwalwasser & Dykema, 1930; Seashore, Lewis, & Saetveit, 1960). However, relatively few standardized tests examine musical abilities acquired, in large part, as a result of instruction.

This scarcity of published materials is reflected in the apparent absence of systematic assessment of music abilities within school music programs. As Zerull (1990) stated:

> The lack of standardized arts curriculum, the subjective nature of the arts and the highly individualized character of arts instruction, and the division among arts educators on the benefits and propriety of testing all contribute to the lack of any widespread, comprehensive, and systematic assessment in the arts. (p. 19)

This view was corroborated by Abeles, Hoffer, and Klotman (1994): "Few music teachers consider assessment as they plan their instruction. When assessment does occur it is likely to be motivated by an administrative requirement like the assigning of grades" (p. 304).

The lack of systematic assessment in music education may be attributed, in part, to an oversight in the requirements of many teacher education programs. This situation, and its subsequent ramifications, was addressed by Colwell (1991) who contended that "evaluation is not a part of most college curricula, so few music educators comprehend the complexity of the discipline and the uses to which evaluation can and should be put" (p. 248). Further, Colwell noted that the lack of understanding with which music educators regard measurement and evaluation may contribute to the view that much of what is taught in music cannot be evaluated.

On the other hand, even though many teachers may disregard procedures for the systematic measurement and evaluation of student performance (Colwell, 1982), "evaluation of student achievement necessarily occurs at all levels of music education" (Radocy, 1989, p. 30). As such, music educators need to know what and how to evaluate (Radocy, 1989). As Boyle and Radocy (1987) explained:

> . . . it is clear that music teachers, administrators, and others necessarily must make and do make decisions that influence learners' instructional opportunities and experiences . . . Not only do instructional and curricular decisions affect individuals, but they also affect broad school, community, and societal goals for music education. It is therefore imperative that such decisions be made from a strong information base. (p. 1)

If indeed music teachers lack a strong information base from which to direct curricular and instructional decisions, it follows that the lack of supportive data may contribute to the decline of music education in favour of other more basic subject areas. In Colwell's (1991) words:

> If music programs have been reduced in order to accommodate more reading or mathematics, this is partly because of evaluation data from reading and math that was convincing. If fiscal constraints have reduced music programs, evaluation data on budgets were used in the fiscal decision-making process. In order for music programs to retain their strength, the supportive data on goals and achievement must be as compelling as the data for the alternatives. (p. 260)

Moreover, as Reimer (1989) contended, "there is no reason why general music should not be at least as expert in using evaluation and measurement as any other basic subject in the schools" (p. 172). In support, Whybrew (1971) and Zerull (1990) concluded that the systematic assessment of student achievement must be addressed if music is to be considered a legitimate school subject.

Many important objectives of school music curricula may be assessed through measurement techniques. This includes cognitive and performance behaviors encompassing abilities in the area of listening and perception (Colwell, 1991). As detailed in Boyle and Radocy (1987), such learnings, when stated in behavioral terms, "are valuable for the test maker because they suggest ways of demonstrating particular knowledge or attitudes that lend themselves to test items" (p. 108). However, it appears that widespread use of so-called tests of music achievement is hindered, in part, by "the lack of consensus among music teachers as to what specific outcomes should be expected as a result of instruction in music" (Lehman, 1968, p. 57). Further, standardized tests assess only a small portion of the objectives addressed in a comprehensive music curriculum (Colwell, 1991).

Not all musical behaviors are easily measured (Abeles, Hoffer, & Klotman, 1994; Lehman, 1993). Further, all objectives referenced to any given music program may not be measured suitably by means of paper-and-pencil test items. In consideration of this, the present study is confined to those objectives which describe student behavior in terms of the ability to read and/or write tonal melodies in traditional Western notation. This aspect of music ability is represented in published tests of music aptitude (Gaston, 1957) and music achievement (Colwell, 1968, 1970a; Farnum, 1969; Gordon, 1971a). However, for the most part, these test constructors examined music reading in terms of the ability to discern differences between music heard aurally and corresponding visual representations—in this case, music written in traditional notation. It is apparent that tests which concentrate solely on this ability under-represent the domain of content (Messick, 1989b). As presented in the context of many school music programs, music reading is a multi-faceted construct. The test items developed and validated in the present study reflect this view.

Problem and Sub-Problems

The problems addressed in the study were to identify the various ways in which music literacy, the ability to read and/or write tonal melodies in traditional Western music notation, is addressed in teaching materials currently employed in school music programs at the grade 5 level and, using this information, to devise pencil-and-paper test items by which students may demonstrate ability in these areas. More specifically, the following questions were addressed:

1. What is the content-related evidence of validity for test items devised to assess student attainment at the grade 5 level in terms of the ability to read and/or write tonal melodies in music notation in relation to selected textbook series, methodologies, and curriculum representative of course content in the upper-elementary grades?
2. What are the administrative requirements for the test items?
3. What are the cognitive strategies employed by students while responding to these test items?
4. What are the indices of difficulty and discrimination for these test items?
5. What is the reliability of these items when viewed as an aggregate?

Delimitations

In the present study, the development and validation of test items was delimited in the following ways:

1. In this study, music was examined in terms of traditional Western notation. Nonstandard representations of melodic and rhythmic stimuli were not considered.
2. In this study, the ability to read and/or write music was explored in terms of Western European tonal music. It did not involve the perception of atonal stimuli.
3. No attempt was made to evaluate success in music reading in terms of the pedagogical philosophies utilized by teachers involved in this study.

The first two delimitations guided the structure of test items. The third delimitation narrowed the focus of the investigation—thus increasing the feasibility of this study.

Definition of Terms

Psychologists and musicians have examined the fundamental nature of understanding music for several decades. However, "despite nearly a century of study by educators and psychologists, there is still much controversy regarding definitions, measurement, and implications of music ability" (Boyle, 1992, p. 247). For this reason, the extended definitions presented below trace the meaning of the terms capacity, ability, aptitude, and achievement from the position of Seashore (1919) to the views of researchers currently working in the field (Boyle, 1992; Radocy & Boyle, 1988). The discussion of each of these definitions concludes with a description of how the term was used in the present study. The discussion of the extended definitions is followed by a series of general definitions for the terms tonality, melody, temporal organization, perception, audiation, music literacy, and notational discrimination.

Extended Definitions

Capacity

Seashore (1919; 1938) viewed capacity as a hierarchy of independent sensory talents—the sense of pitch, the sense of time, the sense of loudness, and the sense of timbre. As Seashore (1938) explained, the basic capacities are elemental "meaning that they are largely inborn and function from early childhood" (p. 3). As an investigator of inheritance, Seashore was not interested in the degree of achievement attained. Rather, in this sense, he was concerned exclusively with the measurement of inborn capacities (Seashore, 1938), the factors of which, for purposes of experimentation (Seashore, 1919), were isolated from elements of training insofar as this was possible (Seashore, 1947). This view was adopted by Seashore's contemporaries. Kwalwasser (1955) made reference to innate capacity as "the possession of inherited structures" (p. 32) which regulates one's potential to succeed. Farnsworth (1969) concurred, adding that capacities are not evidenced directly, but are inferred from "behavioral manifestations such as test scores" (p. 151).

Lundin (1985) criticized Seashore's (1919; 1938) view of capacity because, by his definition, very little was revealed about the nature of music as a psychological construct beyond the assumption that something was inborn. In Lundin's (1985) words: "This inborn 'something' could be a gland, a structure, or a function or psychic potential. The problem of what actually is inborn bothered Seashore very little. To say that capacity is inborn is sufficient without searching further" (p. 205). In consideration of this, Lundin (1985) defined capacity as "a biological potential serving as a framework within which we develop musical actions" (p. 206). This perspective was adopted by Radocy and Boyle (1988) who defined capacity as that "portion of a person's ability which he or she possesses as a result of genetic endowment and maturation" (p. 263). This position was explained by Boyle (1992) who related Lundin's use of the term biological to "maturation variables that might have an effect on musical potential to be considered part of capacity" (p. 250). This interpretation differed somewhat from Seashore (1938) who maintained that inborn capacities do not vary with age, intelligence, or training. However, considering that capacity cannot be measured in the absence of environmental influences, Boyle's (1992) view was adopted by the current author.

Ability

Seashore (1919) began his theory of music perception by distinguishing between capacity and ability:

> The term 'capacity' has reference to the inborn or native power; the term 'ability' is used to designate acquired skill in the use of a capacity. Thus, each of us has a certain native capacity for memory, but we develop various kinds and degrees of ability in the use of capacity. (pp. 14-15)

However, as evidenced in the following test description, Seashore did not continue to make this differentiation: "Each of these six tests purports to measure one of the six capacities or abilities" (1938, p. 383). As used in Seashore (1938), these terms appeared to be interchangeable. This source of confusion was clarified somewhat in recent years. In consideration of this, the current review concentrates on what seems to be an emerging consensus.

Farnsworth (1969) viewed nature and nurture as a joint function. Consequently, he defined ability as the "power to act" (p. 151) but did not specify directly the extent to which this power is dependent upon environmental influences. This perspective was adopted by

Radocy and Boyle (1988) who defined ability as "being 'able' to do something" (p. 262). Radocy and Boyle (1988) did not account for the ways or means by which ability is developed and displayed. However, considering the nature/nurture argument (Farnsworth, 1969), it appears that potential for ability is influenced by several factors including genetics, culture, environment, and education (Boyle, 1992).

Lundin (1985), who conceptualized ability in behavioral terms, used the term acquired skill—"the ability to discriminate different pitches, intensities and intervals, or ability to harmonize melodies, sing at sight, or perform on an instrument" (p. 206). In a similar vein, Boyle (1992) defined ability in terms of testing behaviors. As such, he viewed ability as a "generic term referring to that which students demonstrate on any given music testing task, regardless of the nature of the task" (p. 248). In this respect, ability encompassed a large range of musical behaviors including listening, performing, analyzing, and creating.

It must be noted that, as in the case of Farnsworth (1969) and Boyle and Radocy (1988), both Lundin (1985) and Boyle (1992) did not attempt to unravel the underlying structure of music ability. These questions, a source of discussion and debate for several decades, are beyond the scope of this study. The reader is directed to Shuter-Dyson and Gabriel (1981) for a detailed analysis of factor analytic studies which examined this topic. For purposes of the present study, the meaning associated with the term ability was based on that assigned by Boyle (1992), Boyle and Radocy (1988), Farnsworth (1969), and Lundin (1985). More specifically, musical ability relates to performance in a given task, regardless of the nature of the task or the factors (ie- genetics, culture, environment, education) which contribute to its attainment.

Aptitude

The meaning of the term aptitude has evolved considerably over the past fifty years. It seems, from early reviews (Drake, 1946; Revesz, 1953), that this term is related to

Seashore's (1919; 1938) notion of capacity. Drake (1946) viewed aptitude as the "aptness or quickness in acquiring some skill or ability" (p. 38). He maintained that inherited factors account for this quickness—noting that some people study for several years under expert instruction but are unable to achieve what others acquire in less time and with less effort. Similarly, Revesz (1954) defined aptitude as "that inborn capacity of a person that enables him to realize and develop certain general or specific types of behavior, properties, and capacities" (p. 141). Revesz (1954) further expanded this notion by relating aptitude to potential ability and acknowledging the role played by environmental factors. These aspects of aptitude formed the basis of subsequent definitions.

Lehman (1968) and Radocy (1989) both related aptitude to potential for future achievement. Although these definitions have implications for the appropriate use of test scores obtained from so-called tests of music aptitude, they did not clarify the factors which contribute to behavioral manifestations of the construct. In this respect, a functional definition of aptitude has evolved in the writings of Boyle (1992), Boyle and Radocy (1987), Farnsworth (1969), Gordon (1965; 1971b; 1984), and Radocy and Boyle (1988). According to Farnsworth (1969) music aptitude is only partially innate and implies potential for achievement prior to formal training. This was clarified by Gordon (1965) who described aptitude as a "product of environmental influences and inherited potential" (p. 1). It must be emphasized that, in this sense, potential is not reliant upon innate capacity in the absence of environmental influences. Rather, this potential refers to abilities one may possess as a result of genetic endowment (Seashore, 1919), maturation (Boyle, 1992; Radocy & Boyle, 1988), and abilities which may develop through informal interactions within a musical culture but excluding those abilities whose development depends, for the most part, upon formal instruction (Boyle, 1992; Boyle & Radocy, 1987; Gordon, 1984; Radocy & Boyle, 1988). This was the definition adopted for use in the present study.

Achievement

To Kwalwasser (1955) the term music achievement referred to "accomplishment, attainment—that which one has acquired" (p. 32). Similarly, Gordon (1979b) defined music achievement as "what one has learned" (p. 3). Neither of these descriptions made reference to the types of behaviors involved or the sources of the acquired learning. To this end, achievement was defined more extensively by Boyle (1992), Boyle and Radocy (1987), and Radocy (1989).

Radocy (1989) viewed music achievement in terms of "what a student has done as a result of a relatively specific set of educational experiences" (p. 30). This includes specific task-oriented skills such as "general musical knowledge, knowledge of notation, aural-visual skills, aural skills, and composition as well as performance" (Boyle and Radocy, 1987, p. 157). As observed in Boyle (1992), music achievement encompasses a wide range of accomplishments, many of which are acquired over varying periods of time. It must be noted that, for the most part, musical achievement refers to accomplishments acquired as a result of formal educational experiences (Boyle, 1992; Boyle & Radocy, 1987; Colwell, 1970b; Gordon, 1965; Radocy, 1989). This was the definition adopted in the present study.

General Definitions

Tonality

Tonality refers to a system of tonal relations (Butler, 1983). Within the context of Western tonal music, the system of tonality is based on the diatonic scale. Notes outside the tonal context of any particular scale are referred to as *nondiatonic* tones.

Much of the song material taught in elementary school music programs is written in a five-tone pentatonic scale. The *do pentatonic* scale uses the following configuration: M2, M2, m3, M2, m3.[1] This is distinguished from the *la pentatonic* scale in which the arrangement of intervals is m3, M2, M2, m3, M2.

Melody

Melody refers to "any group of tones heard as a coherent succession" (Piston & DeVoto, 1987, p. 90). *Melodic motion* refers to the upward or downward direction of a melody (Kennedy, 1980). *Conjunct motion* refers to notes which proceed "upward or downward through successive scale degrees" (Mathews & Liebling, 1896/1973, p. 83). *Disjunct motion* refers to notes which proceed upward or downward through successive leaps in the melody line.

Temporal Organization

Rhythm refers to "a system of interconnected temporal relations" (Mainwaring, 1931, p. 185). It is comprised of three basic organizational forms: *beat, meter, and rhythm* (Cooper & Meyer, 1960). *Beat* refers to a succession of equally spaced temporal units, commonly referred to as the pulse. *Meter* refers to predictable groupings of temporal units derived from repeated patterns of accented and unaccented beats. *Rhythm* refers to durational patterns which either divide or extend the beat.

Audiation

Audiation (Gordon, 1984) refers to the perception of musical material when the sound is not physically present. Notational audiation occurs when one audiates music seen in notation. Audiation is often referred to in the literature as *inner-hearing.*

Music Literacy

Music literacy refers to the ability to read and/or write music using the signs and symbols commonly employed in traditional Western notation.

---

[1] M2- major second; m3- minor third.

Notational Discrimination

Notational discrimination refers to the ability to differentiate between the aural presentation of a melody and a corresponding visual representation in which errors have been inserted.

Organization of the Thesis

This thesis contains nine chapters. An overview of the study is presented in Chapter 1. Related literature is reviewed in Chapter 2. The development of the initial test items is described in Chapter 3. This includes a summary of the analysis of selected educational literature which lead to the identification of instructional objectives to which items were referenced as well as an analysis of questionnaire responses obtained from elementary school music specialists who were requested to take the test and evaluate the relevance and representativeness of the items with respect to the specified content domain.

The manner in which students responded to these items was examined by means of six field studies, each of which employed a revised version of the test. Two underlying issues are addressed in these investigations: first, the function of the test and its constituent items as revealed in classical item analysis (Field Studies 1, 3, 5, and 6) and second, the strategies employed by students while solving these items as demonstrated in data obtained through the analysis of student protocols (Field Studies 2 and 4). Analysis of the data obtained in these studies is presented in Chapters 4, 5, 6, 7, and 8. The findings, conclusions, and implications of these test administrations are discussed in Chapter 9.

CHAPTER 2

RELATED LITERATURE

Overview

In 1937, James Mursell published The psychology of music. This was followed closely, in 1938, by Carl Seashore's Psychology of music. Mursell summarized the advance of literature in the field of music psychology while Seashore examined music perception within a psychoacoustical framework. More recently, similar texts have been published by Butler (1992), Deutsch (1982), Dowling and Harwood (1986), Hargreaves (1986), Radocy and Boyle (1988), and Sloboda (1985). More limited in scope, the review that follows concentrates on research which provides a foundation for the present study.

Aural Perception Within a Tonal Context

According to Mursell (1937), "our natural starting point is with melody, for melody is the primary musical phenomenon. Everything in music turns upon it" (p. 99). More specifically, as Sloboda (1985) stated:

> Melody, thus, represents the level of the greatest differentiation in music, the level at which our evaluative and critical faculties are most immediately engaged. It is the aspect of music which is nearest to the 'surface', and that which, for most listeners, most immediately characterizes the music. (p. 52)

The importance assigned to tonal perception is reflected in the number of studies conducted in this area. The psychoacoustical studies reviewed by Butler (1992) examined aural perception in terms of musical sound including experimentation with frequency, duration, and intensity. Although these investigations are useful in delimiting auditory dimensions, "the reduced stimuli employed may be insufficient to engage other perceptual or cognitive processes normally operating during listening to actual music" (Krumhansl, 1983, p. 31). Therefore, during the past two decades, music psychologists have adopted cognitive-musical (Shepard, 1981) approaches to examine melodic perception within a tonal framework (Krumhansl, 1991). Various proposals, summarized by Shepard (1982) and Krumhansl (1991), described these tonal relationships in the form of geometric models. With relation to these models, the present review is limited to Krumhansl's and Shepard's investigations (Krumhansl, 1979; Krumhansl & Shepard, 1979) which reflect widely accepted principles of music theory (Piston & DeVoto, 1987) and provide a foundational framework for the current investigation.

Krumhansl's (1979) and Krumhansl and Shepard's (1979) analysis of the tonal hierarchy was based on the results of an experimental technique known as probe-tone. In their studies, adult listeners rated how well each tone of the twelve-note chromatic scale, the probe tones, completed an ascending or descending C-major scale. They demonstrated that subjects with musical training perceive the scalar structure in relation to a complex hierarchy of tonal functions, whereas subjects without musical training rate melodies in terms of the distance between the probe-tone and the last tone of the scalar sequence. Psychologically, it appears that tonal relationships proceed from the tonic tone, which functions as a stable reference point (Krumhansl, 1979; Krumhansl & Shepard, 1979; Taylor, 1976) to the third (mediant) and fifth (dominant) scale degrees (Krumhansl & Shepard, 1979) which, together with the tonic, form the major triad.[1] These notes, the strongest indicators of tonality, are followed in perceived stability by the remaining diatonic tones. Nondiatonic notes, least related to the tonality, serve to complete the twelve-tone chromatic scale (Krumhansl, 1979). Tones least related to the tonality are the least stable (Krumhansl, 1979). Therefore, whereas diatonic tones move towards notes comprising the tonic triad, nondiatonic tones tend to move toward members of the diatonic scale

---

[1] The reader is advised that Krumhansl (1979) and Krumhansl and Shepard (1979), in referring to the major triad chord, are delineating the major tonic triad chord. In the major scale, major triads are formed on the tonic, subdominant (fourth), and dominant (fifth) scale degrees. However, only the major tonic triad chord employs the tonic, mediant, and dominant tones which form the perceptual framework for tonality as demonstrated in these studies.

(Krumhansl, 1979). In addition, Krumhansl (1979) demonstrated that memory for diatonic tones is greater than memory for nondiatonic tones.

Although Krumhansl's tonal hierarchy provides a psychological basis for the relationships among pitches in a musical context, it does not explain how melodic stimuli are encoded and recalled in memory. According to Dowling (1991), adult listeners hear melodies as integrated wholes in which recognition is influenced by interactions between contour and tonality. As such, he continues:

> . . . we must conclude that the tonal configuration of the melody—the degree to which the pitches match or mismatch the patterns of tonal invariants in the listener's experience—affects the way in which the contour of the melody is heard and encoded. (p. 313)

Melodic memory, in terms of the effects of contour and interval, was examined in a series of transposition tasks in which adult listeners compared given melodic sequences to exact transpositions, same contour lures, and new contour lures (Bartlett & Dowling, 1980; Croonen, 1994; Croonen & Kop, 1989; Cuddy, Cohen, & Mewhort, 1981; Dowling, 1991; Edworthy, 1985). Melodies were identified as tonally strong, tonally weak, or atonal (Bartlett & Dowling, 1980, 1988; Cuddy, Cohen, & Mewhort, 1981; Croonen, 1994; Croonen & Kop, 1989; Cross, Howell, & West, 1983; Cuddy, Cohen, & Miller, 1979; Dowling, 1991; Edworthy, 1985; Long, 1977; Taylor 1976).

Before referring specifically to studies in this area, it is necessary to explain how the terms "exact transpositions," "same contour lures," and "new contour lures," are used in the psychological literature (Bartlett & Dowling, 1980). In exact transpositions a standard melody is followed by a comparison melody which maintains both the interval structure and contour of the original. In this case, the comparison melody is composed in a different tonality than the original. Lures are comparisons which differ from the standard melody in terms of interval structure and/or contour. More specifically, same contour lures maintain the contour of the original standard amid manipulations of the interval. New contour lures involve transformations of both contour and interval structure. The terms manipulation and transformation refer to alterations applied to tonal lures.

The ease with which adult listeners are able to compare standard sequences to transpositions is related, in part, to the tonal structure of both the standard sequence and the comparison melodies. For example, Dowling (1978) illustrated that exact transpositions and same-contour lures are distinguished easily from atonal sequences. Dowling (1991) suggested that the more closely related the tonality of the standard and same-contour lures, the more likely they are to be judged as similar. Cuddy, Cohen, and Miller (1979), and Cuddy, Cohen, and Mewhort (1981) maintained that this similarity is due, in part, to the diatonic structure of the transposed sequence. They demonstrated that transpositions to near keys, which, for the most part, preserve the diatonic sequence of the standard melody, are more easily remembered than transpositions to keys with conflicting diatonic structures. Bartlett and Dowling (1988) demonstrated that pairs of scalar and nonscalar melodies are regarded as less similar when the more scalar melody is heard first. In this case, it appears that the first melody establishes a set of assumptions which the second melody violates, thus inducing the perceived dissimilarity between the sequences. Dowling (1991) showed that when the salience of key information recedes, due to a time delay in which adult listeners hear melodies in other tonalities, subjects base similarity judgments not on contour, but on more detailed interval information. As such, Dowling (1991) concluded that melodic information is perceived as an integrated whole guided by aspects of contour, interval, and tonality.

Although tonality influences the perception of contour and interval (Dowling, 1991), the storage and retrieval of melodic sequences is also dependent on factors such as musical training (Dowling, 1978, 1986) and sequence length (Edworthy, 1985). Dowling (1978, 1982) and Edworthy (1985) illustrated that, in both short-term and long-term memory, contour is retained separately from exact interval size. In terms of adult listeners, both musicians and nonmusicians are able to accurately identify same-contour comparisons

(Dowling, 1978). However, Dowling (1986) reported that musically trained adults employ either contour or interval information in long-term recognition tasks. Edworthy (1985) demonstrated that for trained musicians, interval information, while not as salient as contour in sequences of up to 11 notes, is more reliable in sequences comprised of 15 notes. Furthermore, perceptions of interval and contour are equally reliable in sequences of approximately 13 notes (Edworthy, 1985). In summarizing the results, Edworthy concluded that, "on transposition, contour information is immediately precise but is lost as melody length increases. Interval information is initially less precise but is more resistant to forgetting in longer melodies" (p. 375). Similarly, Croonen (1994) observed that, for adult listeners, tonal series comprised of 7 tones are recognized better than similar tonal configurations comprised of 10 notes.

As noted above, Edworthy (1985) suggested that longer sequences are remembered in terms of exact interval sizes, rather than contour. Similar findings relate to retention time. Dowling (1978) developed a model intended to illustrate how melodies are stored in both long-term and short-term memory. He reported that, for adult listeners, scale and contour are salient in short term (2 second) recognition tasks. Exact interval sizes, however, cannot be deduced within this brief time frame. Supplementary to this, Bartlett and Dowling (1980) concluded that the relative importance of interval information increases as the retention interval is lengthened. This was substantiated by Dowling (1991). In contrast, Croonen and Kop (1989) reported that adult listeners remember interval information within retention times of 1 and 8 seconds. Dowling attributed Croonen and Kop's findings to flaws in experimental design. As Dowling (1991) stated:

> . . . pitch intervals in familiar melodies are remembered much better than those of novel melodies heard for the first time. However, Croonen and Kop used the same two melodies repeatedly throughout their experiments, thereby familiarizing subjects with the test materials and obviating a test of the hypothesis. (p. 308)

In rebuttal, Croonen attributed these conflicting findings to differences in experimental strategy. As Croonen (1994) summarized:

> Croonen and Kop used silent retention intervals in order to investigate whether untrained subjects could abstract interval information from a tone series. Dowling used a continuous running memory paradigm, in which standard series, transpositions, and same-contour lures were presented continuously with a delay of 11 sec between the series; the task was to indicate, for each series, whether it had been seen before. (p. 624)

Croonen's comments highlight the importance of interpreting experimental findings within the structure of the study from which they are generated.

The above-mentioned studies examined how melodic discrimination, in terms of interval and contour, are related to musical training, sequence length, and retention time. In addition, Boltz and Jones (1986), Croonen (1994), Croonen and Kop (1989), and Long (1977) examined contour complexity in terms of the number of directional changes. Long (1977) illustrated that, for adult listeners, memory for M-shaped melodies is greater than V-shaped melodies and postulated that the former places greater demands on attention to interval size and absolute pitch. Boltz and Jones (1986) reported that, for adult listeners, sequences with more directional changes are more difficult to remember than sequences with fewer changes. This is in agreement with Morrongiello and Roes (1990), who examined memory for contour in 5- and 9-year-olds. Contrary to these findings, Croonen and Kop (1989) demonstrated that, in terms of moderately experienced adult listeners, the number of contour changes does not influence the ability to discriminate tonal imitations and exact transpositions. They concluded that:

> . . . interval information can be extracted from sequences within a very short period of time and that one of the major dimensions of tone sequences is not tonality or nontonality, but rather the degree of tonal clarity.[2] This characteristic even may be superior to the number of contour changes. (Croonen & Kop, 1989, p. 49)

This view was supported in Croonen (1994).

Psychological research in aural perception has progressed significantly since Krumhansl (1979) and Krumhansl and Shepard (1979) demonstrated the psychological function of individual notes within the tonal hierarchy by means of the probe-tone technique. Subsequent to their study, psychologists have examined how pitches are encoded and recalled in memory. Findings were based, for the most part, on experimental studies which required adult subjects to compare standard melodies to exact transpositions and tonal lures. Bartlett and Dowling (1986), Cuddy, Cohen, and Miller (1979), and Cuddy, Cohen, and Mewhort (1981) demonstrated that standard and comparison melodies composed within closely related diatonic frameworks are more likely to be judged as similar than melodies with conflicting diatonic structures.

---

[2] Tonal Clarity- The degree to which the melodic material conforms to the specifications outlined in the tonal hierarchy described by Krumhansl (1979).

In addition to tonality, melodic perception is influenced by factors such as contour (Boltz & Jones, 1986; Croonen, 1994; Croonen & Kop, 1989; Dowling, 1978, 1982; Edworthy, 1985), musical training (Dowling, 1978, 1986), retention time (Dowling, 1978; Croonen and Kop, 1989), and sequence length (Croonen, 1994; Edworthy, 1985). Of particular note are Croonen (1994) and Croonen and Kop's (1989) studies which demonstrated that memory for pitches depends more upon the degree of tonal clarity than the number of contour changes. Their findings, however, conflict with those of Boltz and Jones (1986) and Long (1977). Therefore, it appears that further research is needed to determine, in more exact terms, the extent to which tonality influences melodic perception with relation to contour.

Tonal And Durational Relationships

The perception of melody has thus far been portrayed as an integrated whole influenced by relationships among contour, interval, and tonality (Dowling, 1991). Additionally, since melody is considered fundamentally as "a series of successive intervals" (Deutsch, 1969, p. 333), it is essential to examine the relationships which may exist between the melodic and temporal aspects of aural perception. According to Jones (1993), "anything that 'marks' beginnings and endings of coherent pitch groups, such as relatively large pitch intervals, must mark also the durations of these groups" (p. 75). Conversely, rhythmic accents strengthen tonality "by attentionally highlighting key specific information or confirming expectancies about upcoming tonal relationships" (Boltz, 1989a, p. 756). This association is defined as Joint Accent Structure (Jones, 1987, 1993).

Before reviewing the research literature in this area, it is necessary to discuss briefly how accents function within a melodic context. Jones (1993) described three types of melodic accents: contour accents which coincide with changes in pitch direction, interval accents which mark large skips in pitch, and tonal accents which correspond to changes in tonality. The majority of temporal accents are associated with variations in the durational value of individual notes and the presence of silent beats termed 'rests' in music (Jones, 1993). As Jones (1993) stated:

> . . . in integrating melody and rhythm we can ask how melodic and temporal accent structure combine to strengthen a common or joint accent structure. We are also interested in how resultant accent groupings of various strengths define or mark different higher order time spans within the whole tune. (p. 78)

By manipulating relationships between melodic and rhythmic accents, Boltz (1989a, 1989b, 1991), Boltz and Jones (1986), Deutsch (1980), and Jones, Boltz, and Kidd (1982) examined how joint accent structure influences melodic perception in adult listeners. Deutsch (1980) reported that musical memory improves when temporal accents coincide with melodic context. On the other hand, patterns are more difficult to perceive when temporal factors cause melodic accents to sound at unexpected intervals (Boltz, 1991; Deutsch, 1980).

Rhythmic context, in relation with joint accent structure, influences higher order melodic relationships at phrase endings. Melodies characterized by the presence of notes from the tonic triad (Boltz, 1991) or enhanced by the leading tone-to-tonic progression (Boltz, 1989b) are judged as most complete when the underlying temporal accent structure corresponds with these tonal functions. The perceived resolution of melodic progression is emphasized further by the presence of the tritone (Boltz, 1989a). In addition, the highest degree of completion is observed for melodies in which the temporal and tonal structures correspond at phrase endings (Boltz, 1986b). More specifically, within this study, lower ratings of resolution were allocated to an *Accent Early* condition, demonstrated by melodies in which the temporal accent structure caused the endings to occur earlier than expected.

While accent patterning is acknowledged, it is possible that under certain conditions adult listeners respond to temporal and tonal variables as distinct entities (Fiske, 1982; Monahan & Carterette, 1985). Fiske (1982) suggested that melodic and rhythmic aspects of complex unknown musical phrases are processed separately on a note-by-note basis. Monahan and Carterette (1985) contended that adult listeners demonstrate a higher salience for durational as opposed to melodic dimensions. In comparison to solo musicians, there is a slight tendency for ensemble players to attend more to the rhythmic aspects of music. Tasks involving the discrimination of rhythmically consonant and dissonant melodies are performed better by musicians than nonmusicians (Monahan, Kendall, & Carterette, 1987).

Contrary to the model of joint accent structure is a theory of independence which contends that melody and rhythm are perceived psychologically as distinct hierarchies (Palmer & Krumhansl, 1987). In Palmer and Krumhansl's (1987) study, adult listeners judged the completeness of three types of phrases, all based on the same theme adapted from a fugue composed by J.S. Bach:

> Segments in the *pitch condition* preserved the original pitch pattern, but all tones had the same duration (equitemporal); segments in the *temporal condition* preserved the original temporal pattern, but all tones had the same pitch (equitonal); and segments in the *melody condition* retained both the original pitch and temporal patterns. (p. 118)

The low correlation reported between pitch and temporal conditions ($r = .27$, $p > .15$) indicated that their effects do not coincide. Therefore, the authors concluded that "melodic phrase judgments in a Bach fugue subject are based on two independent types of structure: temporal and pitch information" (p. 125). Further, Krumhansl (1991) contended that the effects observed in joint accent structure were due, in part, to experimental manipulations contrary to situations occurring in real music:

> To say that they [temporal and tonal structure] are independent (noninteractive) is simply to say that the effects do not require supposing that additional factors emerge when both tonal and rhythmic components are varied. If the two components are varied in a way that is mutually incompatible or inconsistent, then they may well interfere with one another. (p. 298)

Krumhansl's (1991) concerns were addressed by Boltz (1993). In Krumhansl (1991) subjects were presented with unfamiliar folk tunes in which the initial three phrases were followed by a final variation, two phrases in length, in which all notes were of equal rhythm. The presentations were designed to assess whether temporal and melodic aspects of melodic structure influenced temporal expectancies independently or interactively. Boltz (1993) demonstrated that, for adult listeners, perception of temporal structure depends jointly upon temporal and tonal functions—in this case accents and phrase endings. She concluded that anticipatory attending is "guided by temporal accent structure, which takes a listener to particular points in time so that the melodic information that is there can be processed for its underlying tonality and organizational status" (p. 598).

In summary, within the psychological literature, there are two schools of thought regarding the interaction of temporal and tonal stimuli: the theory of joint accent structure and the theory of independence. Experimental findings representing both sides of this issue are presented in the current review. To date, it may be premature to accept the theory of joint accent structure over the theory of independence, or vice versa. At this juncture, it may be sufficient to say that these studies inform our understanding of how listeners perceive the interactions between melodic and rhythmic stimuli. Additionally, however, it must be noted that Boltz's (1993) experimental application of unfamiliar folk material may have satisfied Krumhansl's criticisms, thus influencing future work in this area.

A Developmental Perspective

Musical concept development is dependent upon experience and growth in listening skills (Simons, 1986). As Zimmerman (1985) stated, "through experiences with music, the conceptual framework develops from simple concepts to an interlocking network of operative musical concepts, which are based on relational transformations" (p. 69). Although there are individual differences among studies, it is generally noted that concepts in melody and harmony develop later than those of beat, tempo, timbre, and dynamics (Greenberg, 1976; Hufstader, 1977; Sergeant, 1973). In terms of melody, children initially focus on the contour (Bartlett & Dowling, 1980; Morrongiello, Trehub, Thorpe, & Capodilupo, 1985; Ramsey, 1983). Conceptualizations of the functional structures of tonality emerge during the preschool and early school years (Dowling, 1982, 1988; Krumhansl & Keil, 1982; Simons, 1986).

Several longitudinal studies examined the development of song in young children (Davidson, 1985; Davidson, McKernon, & Gardner, 1981; Jersild & Bienstock, 1934; McKernon, 1979; Moorhead & Pond, 1978; Ramsey, 1983). Moorhead and Pond (1978) described the spontaneous songs of young children as rhythmically free and flexible with a melodic movement predominated by small steps and frequent shifts in tonal center. It appears that children sing small intervals, such as unisons, seconds, and thirds, more accurately than larger intervals, most specifically a seventh (Jersild & Bienstock, 1934; McKernon, 1979; Moorhead & Pond, 1978; Ramsey, 1983). In addition, Ramsey (1983) reported that children sing descending patterns more accurately than ascending patterns.

Gradually, over a 4 year period, these spontaneous melodies are supplemented by renditions of standard tunes which eventually emerge as accurate representations of the original versions (Davidson, 1985; McKernon, 1979). According to Davidson (1985), "the path of development runs from a loose rendering of a song, by means of a single contour type, toward a more specific rendering of a song, as a set of differentiated contours and specific pitches" (p. 34). More specifically, while 2- and 3-year-olds reproduce the approximate contour of melodic phrases, 4-year-olds represent phrases in terms of stable scale patterns (Davidson, McKernon, & Gardner, 1981). Whereas 4-years-olds often sing subsequent phrases in new keys (Ramsey, 1983), by the age of 5 or 6 children maintain a stable tonality throughout the song (Davidson, McKernon, & Gardner, 1981).

The extent to which vocal performance influences pitch discrimination is subject to conflicting accounts. According to Joyner (1969), competent singers are able to discriminate pitches better than poor singers. On the other hand, Apfelstadt (1984), Geringer (1983), and Pederson and Pederson (1970) reported a low correlation between pitch discrimination and vocal pitch production. This coincides with Duell and Anderson (1967), Serafine (1988), and Webster and Schlentrich (1982) who suggested that the ability to discriminate melodic material is influenced by maturation. As Geringer (1983) concluded, vocal performance and pitch discrimination are two separate abilities on which both maturation and training exert an influence. Further, with respect to training, Buckton (1977) reported no difference between instrumental experience or vocal practice and improvement in pitch discrimination.

Andrews and Deihl (1967) and Van Zee (1976) illustrated that preschool childrens' achievement on pitch discrimination tasks is negatively effected by the use of terminology such as 'high/low' or 'up/down'. To this end, Hair (1977), Scott (1979), and Webster and Schlentrich (1982) demonstrated that nonverbal, performance-based response modes may be the best way for young children to demonstrate the ability to discriminate changes in pitch. However, Webster and Schlentrich (1982) noted that, since 34% of the children in their study performed below chance, one should not assume unconditionally that preschoolers conceptualize pitch movement. This is in conflict with Scott (1979) who concluded that preschool children are capable of demonstrating an understanding of pitch register, melodic contour, and interval size. Consequently, in consideration of both Scott (1979) and Webster and Schlentrich (1982), it appears that additional research is needed to examine the extent to which young children are able to perform these tasks.

Pitch discrimination in children is also related to the type of stimuli employed in the task. Petzold (1966) demonstrated that, with respect to singing tasks, children respond less accurately to flute and piano than to violin or soprano voice. Hermanson (1972) reported that children match pitches more accurately when singing with a woman's voice and least accurately with the piano. Small and McCachern (1983) and Yarbrough (1991) observed that children respond more accurately to a female vocal model as opposed to a male vocal model. This is congruent with Green's (1990) finding that children match pitches most accurately with a child's voice, followed by the female and male models, respectively.

Melodic discrimination in children is also influenced by the acquisition of tonality. Dowling (1988) contended that children as young as 3 years of age demonstrate a preference for diatonic as opposed to nondiatonic melodies. By the ages of 4 to 6, children are superior at discriminating alterations of a semitone in a diatonic context as opposed to a nondiatonic context (Trehub, Cohen, Thorpe, & Morrongiello, 1986). Children of elementary school age prefer tonic triad tones over other tones of the diatonic scale (Krumhansl & Keil, 1982).

Above age 9 or 10, children discern intact melodies more easily than random pitches while below the age of 8 children are equally able to discriminate both tonal and atonal material (Morrongiello & Roes, 1990; Serafine, 1988). Serafine (1988) maintained that a turning point in idiomatic discrimination appears at age 8. This is in agreement with Petzold (1966) who noted a significant change between grades 1 and 2 with a plateau no later than grade 3 and Gordon (1979a) who described the stabilization of developmental music aptitude at age 8. Serafine (1988) contended that 10-year-olds perform some recognition tasks significantly better than 11-year-olds. She termed this "the sixth grade slump".

McDonald and Ramsey (1979), Ramsey (1983), and Sergeant and Roche (1973) demonstrated that young children initially perceive melodic contour. This is followed by the conceptualization of more refined intervallic relationships. Bartlett and Dowling (1980) reported that, below the age of 8, children are unable to differentiate between contour-preserving transformations in which intervallic relationships are manipulated and direct transpositions in which both intervallic relationships and contour are preserved. In a similar study Morrongiello, Trehub, Thorpe, and Capodilupo (1985) demonstrated that children ages 4 to 6 detect both contour preserving and contour violating transformations. However, they noted that recognition is superior in sequences that incorporate changes in contour as opposed to those in which the contour is maintained. They reported that "the cue of contour change did not appear to operate in an all-or-none fashion: greater degrees of contour change (five versus two intervals) were associated with greater discriminability" (p. 289). In addition, Morrongiello, Trehub, Thorpe, and Capodilupo (1985) demonstrated that discrimination improves when tones are presented at a faster rate of 4.5 tones/sec. as opposed to a rate of 2.5 tones/sec. Further, an increase in the rate of presentation appears to improve the discrimination of tasks in which contour is transformed to a greater degree than those tasks in which contour is maintained. This superior performance on contour manipulated comparisons implies that melodic perception in children is dependent upon contour rather than interval. The comparatively poor performance of young children in Bartlett and Dowling's (1980) study was attributed to differences in experimental procedures (Morrongiello, Trehub, Thorpe, & Capodilupo, 1985). More specifically, in the former study presentation of the standard melody was limited to a single performance, while in the latter experiment children heard each melody several times.

In summary, it appears that music skills are acquired developmentally with concepts in melody and harmony developing later than those of beat, tempo, timbre, and dynamics (Greenberg, 1970; Hufstader, 1977; Sergeant, 1973). Studies reveal that the spontaneous songs of young children gradually evolve into accurate portrayals of standard tunes. The extent to which the ability to discriminate tonal stimuli is related to the ability to sing in-tune is subject to conflicting accounts (Apfelstadt, 1984; Geringer, 1983; Joyner, 1969; Pederson and Pederson, 1970). There is evidence, however, to suggest that children as young as 3 years of age respond to music composed within the tonal idiom (Dowling, 1988; Trehub, Cohen, Thorpe, & Morrongiello, 1986). Moreover, beyond the age of 8, it appears that children are better able to discern intact melodies as opposed to random pitch sequences (Serafine, 1988). This corresponds to a plateau in musical development described by Petzold (1966), and Gordon (1979a).

Music Reading

Sloboda (1978) observed that the ability to read and write music notation is "an irreplaceable asset to anyone who indulges in musical activity" (p. 3). In view of this, it is surprising to note that few research studies investigate the psychological foundations of music reading. As Hodges (1992) observed: "Explicit theories of music reading, theories that would organize knowledge and research about music reading into a system of assumptions, principles, and procedures do not exist" (p. 469). The current review is limited first, to psychological studies which examine music reading in terms of visual perception and, second, to educational studies which, for the most part, examine music reading in terms of score manipulation.

Psychological Studies

Hodges (1992) described the mechanics of music reading as involving:

> . . . a rapid series of stops (fixations) and starts as the eye focuses on pertinent information and then sweeps (saccade) to the next focal point. Information is brought into the visual system at a fixation, when the eye is not moving and is focusing on a circular area about one inch in diameter. (p. 466)

One of the first to investigate this phenomenon was Bean (1938). In his study, pianists were required to view musical excerpts briefly and subsequently to perform them from memory. Bean (1938) concluded that proficient sight readers memorize groups of musical patterns rather than individual notes. However, as Sloboda observed, since subjects were required to play the piano, it is not clear "how much of the superiority was due to perceptual factors and how much due to factors of manual response" (1976b, p. 3). Nonetheless, this research is significant in that it anticipated more recent investigations in this area, most notably the series of studies conducted by Sloboda (1974, 1976a, 1976b, 1977, 1978).

Sloboda (1974, 1977) examined notational recall in relation to what he terms the eye-hand span.[3] In these studies, musicians were shown photographic slides of keyboard music. They were instructed to begin playing a piece as soon as the music was displayed and, after the slide was switched off, continue to play until they could not remember the notes. Sloboda (1974) reported that pianists with a eye-hand span of three to four notes make more mistakes than pianists with eye-hand spans of six to seven notes. In a similar study, Sloboda (1977) observed that musicians have a greater eye-hand span when reading tonal music as opposed to random atonal sequences. Further, Sloboda (1977) demonstrated that the eye-hand span tends to extend to phrase boundaries, suggesting "that knowledge of abstract structure is of importance in the organization of immediate visual processing of text" (Sloboda, 1977, p. 117).

These findings (Sloboda, 1977) supported his earlier research (Sloboda, 1976a). In Sloboda (1976a), pianists were required to sightread four pieces of keyboard music of the Baroque-Classical periods in which notational errors were introduced "by raising or lowering the pitch of a note by one position on the stave" (p. 23). Sloboda (1976a) noted that errors were detected less often when they occurred in the middle of phrases than when they occurred at the beginning or ending of phrases. Sloboda (1976a) concluded that, with reference to error detection in the middle of a phrase, subjects use aspects of the structure to help them make inferences.

As summarized above, notational recall of keyboard music is examined in Sloboda 1974, 1976a, and 1977. In contrast, notational recall of random pitch sequences was examined in Sloboda (1976b). Findings in this study indicated that, after a 2 second viewing, musicians reproduce sequences of four or more notes in length better than nonmusicians. However, for exposure times of 20 ms, musicians do not perform significantly better than nonmusicians. Sloboda (1976b) hypothesized that:

---

[3] Eye-hand span- the number of notes that are played after the music is no longer visible.

> . . . these results might be accounted for by proposing that two different coding processes were occurring. 'Visual' coding was held to be a rapid process which filled a small capacity store in about 100 ms. 'Name' coding was held to be a slow process, generating one more permanently stored item every 100 ms or so. The results would then suggest that masters had superior 'name' codes than did novices but were not better at 'visual' coding. (Sloboda, 1978, p. 323)

This is documented further in Sloboda (1978). In this study musicians and nonmusicians were presented with standard staves on which noteheads were printed in four types of stimulus configuration:

> Type 1 configurations were 'straight-line' contours obtained by placing notes on the four spaces between stave lines in ascending or descending sequence. Type 2 configurations had one major direction change, thus resembling a rotated 'L' . . . Type 3 configurations had two major direction changes, thus resembling a rotated 'Z'. . .Type 4 configurations had no distinctive pattern, but were used as 'dummy' trials. (p. 326)

Subjects were required to reproduce the notation from memory under exposure times of 50 ms and 500 ms. Findings indicate that musicians do not perform better than nonmusicians under an exact-accuracy measure. However, musicians perform significantly better in terms of absolute contour, illustrating that musicians retain more information about the relative position of the notes than their untrained counterparts.

As observed in Sloboda (1976b, 1978) musicians, as compared to nonmusicians, demonstrate superior ability in the notational recall of random atonal sequences. In a related study Halpern and Bower (1982) examined similar performance in terms of musically consonant and dissonant sequences. They reported that musicians remember the visual patterns of consonant melodies better than those of dissonant melodies. Nonmusicians, on the other hand, show no preference. According to Halpern and Bower (1982), this is due partly to the musicians' ability to recognize patterns commonly associated with tonal music.

The visual perception of tonal material is examined further in Goolsby (1989). In this study subjects sang four melodies, three times each, with a 4 minute practise session separating the second and third performances. Eye movements were monitored with computerized equipment which recorded eye positions 1,000 times per second. Goolsby (1989) demonstrated that notation in which the notes are closely spaced is read with fewer and shorter fixations than notation in which the notes are spaced further apart.

In summary, it appears that reading ability is related to the tonal and visual qualities of the music being read. First, musicians as opposed to nonmusicians are better able to recall both atonal (Sloboda, 1976b, 1978) and tonal sequences (Halpern and Bower, 1982). Second, as reported in Goolsby (1989), closely spaced notes are more accurately read than notes spaced further apart. Third, musicians are able to retain contour information more efficiently than their untrained counterparts. Fourth, in trained musicians sightreading is guided by structural information, such as phrase endings (Sloboda, 1976a).

Educational Studies

As noted in Hodges (1992), educational studies in the area of music reading "are so scattered as to render overall conclusions exceedingly difficult" (p. 467). An exception to this was Petzold's (1960) seminal study in which music reading ability was examined in terms of vocal responses to notational stimuli. In this study, children in grades 4, 5, and 6 were presented with tonal sequences comprised of from three to six notes. For each item, children were instructed to sing a melodic sequence presented in traditional Western notation. This was followed by the performance of this melody, played on the piano, after which children were instructed to sing what they heard while looking at the notation. As Petzold (1960) explained:

> Considerable emphasis was given to the idea that hearing an item after the subject had tried to read it was designed to make it possible for the subject to recall the sound of the item the next time it was presented visually, as well as to provide a model against which he could compare the original attempt. (p. 284)

Petzold (1960) reported no difference between boys and girls. Findings also indicate that children read music notation slowly, one note at a time. In concurrence with Bean (1938), Petzold noted that children frequently read the general outline of the structure, but are unable to sing the absolute intervals.

Children's ability to read music notation under the manipulation of certain visual aspects of the printed score was examined by Gregory (1972), Franklin (1977), and Rogers (1991). Gregory (1972) presented grade 7 students with musical scores which employed four different notational styles and, using these scores, compared the students' sight-reading performances in terms of the number of errors. Findings indicated that students do not perform significantly better when presented with traditional notation as opposed to notation in which individual notes are spaced in proportion to their rhythmic duration. While some students preferred the conventional notation in which the beats were indicated above the notes, several students were confused by this notational style. Further, students were unable to cognitively process notation in which stemless note heads were elongated in proportion to their rhythmic duration. Gregory (1972) concluded that instructional programs employing different notations appear useful. However, he cautioned that further research was needed before the implementation of such approaches.

Franklin (1977) experimented with notation in which the song text was written on the staff in place of the notes (termed text notation). Untrained pupils performed better by practising with text notation as opposed to traditional notation with the text printed below the staff. Franklin (1977) concluded that results are poorer with traditional notation because attention is divided between the notes and the text.

Rogers (1991) examined the extent to which the use of colour-coded notation influenced fifth grade and sixth grade students ability to perform music from memory, sight-read, and identify notes by letter name. Findings indicated that coloured notation positively influences students sight-reading ability, but does not effect their ability to perform familiar music. In addition, there was a positive affective influence in that 65% of the students expressed a preference for the coloured notation. However, affective benefits aside, one wonders the extent to which the use of colour-coded notation encouraged students to read notes by colour and not by association with their placement on the staff.

The efficacy of computer assisted instruction (Willett & Netusil, 1989) may indicate a future direction for research in this area. In this study, two groups of fourth grade students were introduced to the notes in the bass clef. The first group received instruction by traditional means: naming notes on the chalkboard and drawing notes on slates. The second group received similar drills by means of computer assisted instruction. Willett and Netusil (1989) reported that, within a pretest-posttest experimental design, the computer group performed significantly better than the traditional group. As such, they concluded that the computer is an effective means by which to practise note reading.

To date, there does not exist a theory of music reading to explain how children visually interpret music notation. Petzold (1960) reported that children read notation one note at a time, attending to the overall contour rather than the exact placement of the notes. No educational studies expand upon his work. Gregory (1972), Franklin (1977), and Rogers (1991) examined music reading under conditions of score manipulation. Finally, Willett and Netusil (1989) examined the effects of computer assisted instruction on reading achievement. With this investigation, they may have identified a direction for future studies in this area.

Standardized Music Tests Incorporating Tasks of Music Reading

Psychologists and musicians have pondered the fundamental nature of music understanding for several decades. This concern lead to the development and publication of numerous music tests in the areas of achievement, aptitude, attitude, and performance—many of which were reviewed extensively by Boyle and Radocy (1987), Colwell (1970b), George (1980), Lehman (1968), Lundin (1985), and Shuter-Dyson and Gabriel (1981). The current review is limited to prominent standardized tests which, in part, examine student performance in terms of the ability to read and/or write traditional music notation. Particular attention is paid to the structure of items which assess notational discrimination.

Test of Musicality (Gaston, 1942, rev. ed. 1950, 1956. 1957)

The purpose of Gaston's Test of Musicality is to provide an assessment of music ability and a profile of general interest in music. General interest in music is assessed in the form of a questionnaire (the first 17 items) with questions such as "Is a phonograph played in your home?".

Gaston believed that musicality is related to an awareness of the tonal relations within the diatonic scale. This philosophy is reflected in the structure of the test questions which assess music perception. In the first five musical items students hear a single tone followed by a chord and are required to indicate whether the single tone is included in the chord. This is followed by five items in which students indicate whether the aural presentation of a melody is the same as the given notation and, if different, whether notational errors are introduced in the melody or the rhythm. In the next five items students hear an incomplete melody and indicate whether the final note should be higher or lower than the last tone heard. Finally, students are presented with seven items in which a given melody is heard from two to six times. Students indicate whether the repetition is the same as the original or whether differences are rhythmic or tonal. The total test is administered in 40 minutes.

The Test of Musicality was standardized separately for boys and girls for ages 9, 10 to 11, 12 to 13, 14 to 15, and 16 to 18 and correspondingly, for grades 4, 5 to 6, 7 to 8, 9 to 10, and 11 to 12 based on a norming sample of 4,840 students in 26 schools in seven different states. Separate norms were reported for the interest inventory. Gaston reported split-half reliability coefficients of .88 for 653 fourth, fifth, and sixth grade students and .88 for 166 seventh, eighth and ninth grade students. Split-half reliability was .90 for 153 students in grades 10, 11, and 12.

Overall, the Gaston Test of Musicality is easily administered and scored. Each group of tonal questions is preceded by clear directions and two or three practise questions. One wonders, however, whether a test of only 22 items can adequately provide a measure of musically.

Farnum Test of Music Notation (Farnum, 1969; subtest 2 originally published in 1953)

Farnum specified four factors for identifying exceptional music students—1. eye and hand coordination, 2. recognition of music notation, 3. musical memory (tonal patterns), and 4. tonal movement (cadence). The test battery, designed for students in grades 4 through 9, has a total administration time of 40 minutes.

Further analysis of the notation subtest is pertinent to the current study. With a structure similar to that of the Gaston test, students identify the measure in which the aural presentation of a melody differs from that of the given notation. Larson (1959) noted that, of the 40 items, 33 had errors in the melody, 6 had errors in the rhythm, and 1 had errors in both the rhythm and the melody. Bean (1959) observed that items encompass a range of difficulty which allows for discrimination between high and low achievers. Some of the errors in notation involve several notes over several scale steps while other errors involve the alteration of one note by one scale step. Farnum had originally planned to order the items in terms of difficulty. This idea was abandoned in order to separate melodies of similar content (Farnum, 1969).

Criterion-related evidence of validity was demonstrated by correlating scores obtained with the Farnum music test with those obtained on the Watkins-Farnum Performance Scale. Subjects were 152 beginning band students tested after one year of instrumental instruction. Farnum (1969) reported a product moment correlation coefficient of .63 and concluded that the Farnum Music Test satisfactorily predicts scores on the Watkins-Farnum Scale. This claim was challenged by Larson (1959), who viewed the Watkins-Farnum Scale as an unsuitable criterion due to its limited application as a measure of achievement. Reliability coefficients for the total test are not included in the manual. However, as reported in Larson (1959) and Shuter-Dyson and Gabriel (1981), split-half reliabilities for the notation subtest ranged from .78 to .91.

Overall, the Farnum Music Test is easy to administer and score. Each subtest begins with clear directions and three or four sample questions. Each item begins with a vocal announcement of the item number which is followed immediately by the melodies played on the piano. With reference to the cadence test, the short five-note melodies might be heard more easily if a slight pause separated the announcement of the item from the performance of the melody.

Music Achievement Tests (Colwell, 1968, 1970a)

Richard Colwell's Music Achievement Tests are intended to provide a measurement of student achievement in terms of the most important objectives in music education. Designed for grades 3 to 12, the battery is composed of four tests each of which is comprised of three or four subtests. Test 1 (18 minutes) measures pitch, interval, and meter discrimination; Test 2 (28 minutes) measures auditory/visual discrimination as well as recognition of major/minor and tonal center; Test 3 (32 minutes) measures tonal memory and recognition of melody, pitch, and instrument timbre; Test 4 (38 minutes) measures auditory/visual discrimination and recognition of musical style, chord, and cadence. Melodies are played on the piano, violin, viola, and cello. Additional instruments are used for the Instrument Recognition subtest.

In the second subtest, auditory/visual discrimination is divided into two sections. In the first section, students listen to 12 four-measure melodies and identify every measure in which the pitches played in the aural presentation differ from the pitches printed in the given notation. This is followed by 12 four-measure melodies in which students identify every measure in which the rhythm played in the aural presentation differs from the rhythm printed in the given notation. The melodies composed for this test are comprised of half notes, quarter notes, and eighth notes. A more difficult version of the rhythm task (melodies comprised of eighth notes, sixteenth notes, and dotted rhythms) is included in subtest four. As designed for the MAT, items of auditory/visual discrimination may be easier than those of Farnum because, in any given question, students are instructed to attend to either pitch or rhythm as opposed to both pitch and rhythm as required in the earlier test.

Standardization was based on a sample of 20,000 students for Tests 1 and 2 and 9,000 students for Tests 3 and 4. Norms, including raw score, standard score, and percentile score were provided for individual and combined grades (Colwell, 1970b; Lehman, 1972). Criterion-related evidence of validity was demonstrated through the use of teacher ratings. Colwell (1970b) reported a correlation of .92 between scores obtained on the MAT and ratings based on the top and bottom 20% of the class. Reliabilities for the four tests, obtained by the KR 21 method, were .88, .94, .90, and .90 (Colwell, 1970b). It should be noted, however, that this method of computation assumes equal difficulty among items (Crocker & Algina, 1986). Therefore, it represents a lower-bound estimate. The split-half method yielded a reliability estimate of .94 for subtests 1 and 2.

Colwell's Music Achievement Test is easy to administer and score. The four subtests which comprise the battery provide a standardized assessment of certain aspects of music perception—aural skills in particular. Although many aspects of this test are assessed in tests of music aptitude, the innovative subtest on instrument recognition may be of particular interest to educators.

Iowa Tests of Music Literacy (Gordon, 1971)

Edwin Gordon (1971a) constructed the Iowa Test of Music Literacy to provide a measure of music achievement in terms of tonal and rhythmic concepts. It is comprised of six levels, each of which is sub-divided into a rhythmic and a tonal component. Each component is divided further into three subtests: aural perception, reading recognition, and notational understanding. The first three levels are intended for grades 4 to 12 and the following three levels for grades 7 to 12.

In the tonal section of the aural perception subtests, students identify given melodies as major, minor, or tonal. The aural perception of rhythmic material involves identification of meter—duple, triple, or mixed. In the tonal section of reading recognition, students indicate whether the aural presentation of a melody is the same or different than the pitches printed in the given notation. Aural-visual discrimination of rhythmic material follows a similar format. Subtests of notational understanding require students to complete music notation. In terms of melody, melodies comprised of nine notes are played three times. The first four notes are notated and students must complete the pattern by selecting notes from between two alternatives. Levels four and five involve the bass clef and level six employs chord symbols. Similarly, students complete rhythmic patterns with notes, flags, beams, rests, and ties. As Boyle and Radocy (1987) observed, tonal patterns and rhythmic configurations become more complex with each succeeding level.

Percentile ranks for each level were based on a norming sample of 18,680 students (Shuter-Dyson & Gabriel, 1981). Split-half reliability coefficients, based on composite scores for each level, ranged from .87 to .94. Criterion-related evidence of validity was demonstrated by Young (1976) who reported mostly moderate to high intercorrelations among scores obtained on the ITML and the MAT.

Gordon viewed music reading as more than aural-visual discrimination. The notational understanding subtests are an innovative means of measuring notational understanding and, as such, are a valuable addition to the literature. As Lehman (1978) noted, in the ITML battery Gordon has created "the finest tests available for grades 4-12 of these fundamental and important abilities" (p. 183).

Summary

Preference for any particular instrument for measuring student achievement ultimately depends on the educational background of the students. From a review of these tests it appears that aural/visual discrimination is an important aspect of music reading. On the other hand, it is possible that test constructors emphasize this aspect of music reading because it is adaptable to the overall structure of pencil-and-paper music tests. All tests employ melodies in which tonal and rhythmic aspects of music are interrelated. However, it is interesting to note that, in both the Colwell and the Gordon tests, students are instructed to attend to either the melody or the rhythm. The Farnum test, while emphasizing melody, does not make this differentiation. One wonders if this influences the students' ability to identify the notational errors. This is a question deserving of further attention.

Selected Published Music Tests and Gender Bias

The possibility of gender bias in tests has been raised in recent years (Boyle & Radocy, 1987; Linn & Gronlund, 1995). Criticisms center on the use of masculine pronouns and the portrayal of women in traditional roles within the home and the workplace. This concern was illustrated by Tohidi, Steinkamp, and Maehr (1986) who, in an analysis of 70 studies in which the performance of females was compared to that of males on tests of cognitive functioning, found small but consistent gender differences in favour of males. The current review is limited to studies which compare the performance of females and males on selected music tests. Both the claims of test authors and the findings of independent researchers are summarized.

Claims of Test Authors

As portrayed in the Seashore Measures of Musical Talent, Seashore (1919) reported that performance on tasks of discrimination did not vary with gender:

> . . . certain tendencies are shown in groups of records, sometimes in favor of one and other times in favor of the other sex; but, on the whole, it seems certain that such differences except so far as they are due to selection in grouping, may be accounted for as due to the conditions of the test rather than to the sex difference in the psychophysic capacity of pitch discrimination. (p. 56)

This finding was substantiated in the most recent version of the test (Seashore, Lewis, & Saetveit, 1960). As stated in the manual for administration, differences on the basis of gender were small and inconsistent—thus negating the need for separate norms.

Wing (1941) noted that, although boys and girls performed equally on tests of ear acuity, after the age of 14, girls out-performed boys on tests of appreciation. Further examination of this claim was beyond the scope of this study. It must be noted, however, that Wing (1968) later reported that these observed differences were not sufficient to necessitate the preparation of separate norms.

In the Drake Test of Music Aptitude (1957) the test manual reported a slight superiority for females on the test of musical memory. However, the observed difference in performance between females and males was not sufficient to warrant the formulation of separate norms. No significant difference for females and males was reported for the rhythm test. Similarly, in the Music Aptitude Profile (Gordon, 1965) differences in score attainment slightly favoured females. However, Gordon (1965) reported that these differences were educationally insignificant. Gordon maintained this view in both the Primary Measures of Music Audiation (1979b) and the Intermediate Measures of Music Audiation (1982). In addition, this position was adopted by Bentley (1966) who, with reference to his Measures of Musical Abilities, reported no inherent differences in music ability based on gender.

In Colwell's Music Achievement Tests, norms for grades 4, 5, 6, 7, 8, and high school were categorized by geographical location, size or type of school, and type of instruction (music specialist or classroom teacher). Gender was not listed as a significant factor in music achievement. This is contrary to Farnum (1969) who reported that, although gender differences were small, they were not small enough to permit combined norms for females and males. Similarly, separate norms were reported in Gaston (1957). In consideration of available evidence, it is not possible to determine why the findings of Farnum (1969) and Gaston (1957) differed from those of the above-mentioned test authors.

Additional Findings

Gender bias in music tests has been examined in several studies independent of those conducted for purposes of standardization. As the following review demonstrates, findings generally support the claims of the test authors.

Rainbow (1965), in his investigation of the constructs of music aptitude, administered the Pitch Discrimination, Tonal Memory, and Rhythm subtests of Seashores Measures of Musical Talents, the Tonal Memory subtest of Drake's Music Aptitude Tests, and the Kwalwasser-Ruch Test of Musical Accomplishment to 291 students—91 in grades 4, 5, and 6; 112 in grades 7 and 8; and 88 in grades 9 through 12. He reported no significant differences between females and males.

In a longitudinal study, Young (1976) compared student performance on the Tonal Imagery and Rhythm Imagery subtests of Gordon's Music Aptitude Profile, Bentley's Measures of Music Abilities, the second test from Colwell's Music Achievement Tests, and the sixth level of Gordon's Iowa Tests of Music Literacy. He reported that all differences based on gender were insignificant and concluded that "the assumption that the sex of the student has no relationship to score magnitude on an objective music test seems valid" (p. 107). Young's finding concurred with Rowntree (1970) who administered the Bentley Measures of Musical Abilities to approximately 3,000 students between the ages of 7 and 11 and found no difference between the performance of females and males.

McCarthy (1980) examined the effects of individualized instruction of music achievement as measured by the second test of Colwell's Music Achievement Tests. Multiple regression analysis revealed that scores on the achievement measure were not predicted by gender. By similar means Hedden (1982) demonstrated that gender was not a significant variable in predicting music achievement as measured by the first test of Colwell's battery. Conversely, Zdzinski (1992), who examined relationships among selected aspects of parental involvement, music aptitudes as measured by the tonal and rhythmic subtests of Gordon's Music Aptitude Profile (1965), and music achievement as measured by selected subtests of Colwell's Music Achievement Tests (1970), observed that males scored significantly higher on Colwell's tests than females. Zdzinski did not explore the reasons for this observed difference.

Summary

In general, it appears that gender bias is not a concern in tests of music aptitude and music achievement. In consideration of Wing (1941), it is possible that females and males perform differently on tests of music appreciation. However, the further investigation warranted by this claim is beyond the scope of the current study. It is interesting to note that gender differences observed in Farnum (1969) and Gaston (1957) necessitated separate norms for females and males. Considering that, in part, these tests measure students' ability to recognize music notation, this finding may have implications for the current study.

CHAPTER 3

DEVELOPMENT OF THE INITIAL TEST ITEMS

Overview

The development of the initial test items is described in Chapter 3. In the first main section, educational and psychological aspects of test construction are examined. First, selected educational materials which represent course content of upper-elementary music classes were reviewed and analyzed to identify how music reading is represented in the educational literature. Second, selected educational materials were reviewed and analyzed to identify the scalar, metric, and rhythmic structure of songs employed in these sources. Third, psychological studies were reviewed to identify the psychological representation of interval and contour. Information obtained from the latter two reviews formed the basis for the structure of the melodies composed for specific use in item construction. This section concludes with a description of the sequential organization of test items which led to the formulation of the initial test version and a summary of specifications which led to the development of the cassette recording which accompanied items for purposes of test administration.

As will be described in the second main section of Chapter 3, content-related evidence of validity was obtained by means of a questionnaire forwarded to subject-area specialists who were requested to take the test and comment on the representativeness and relevance of the test items as referenced to the content domain identified in the survey of the pedagogical literature (Messick, 1980, 1983, 1989a, 1989b). Further refinement of test items was guided by suggestions for item improvement supplied by these judges.

There was no research literature to guide decisions regarding the time needed to complete the test items developed in this study. Therefore, time allotments were based on the professional judgement of the test author. In order to assess the appropriateness of these time-frames, the test was administered to a grade 6 class within School Division 1.

As detailed in the final section of this chapter, time allotments were adjusted on the basis of these data.

Initial Test Development

Item Construction

Educational Representation of Music Reading

Melodic Characteristics. Content representation involved a review of the following sources in order to identify the characteristics of melody which relate specifically to the aural and visual perception of melodic stimuli: the textbook series Music and You (Staton, Staton, Davidson, Kaplan, & Snyder, 1988; Staton, Staton, Davidson, & Ferguson, 1988; Staton, Staton, Davidson, Davis, & Ferguson, 1988) and Musicanada (Brooks, Kovacs, & Trotter, 1983, 1984a, 1984b); the Kodály method as interpreted in The Kodály Context (Choksy, 1988); the Orff approach as interpreted in Exploring Orff (Steen, 1992); and the Elementary Music Curriculum Guide developed by the Province of Alberta, Canada (1989). Limiting the analysis to those melodic aspects which, in general, are introduced by the end of the fourth grade, 12 characteristics of melody were identified through a review of scope and sequence charts, lesson plans, and pedagogical methods presented in these sources. The 12 characteristics of melody identified are listed in Table 1.

The distribution of melodic characteristics among sources considered is summarized in Table 2. As shown, melodic characteristics 1, 2, 3, 8, 10, 11, and 12 are addressed by the fourth grade in all the materials consulted. Melodic characteristic 6 is inferred in Exploring Orff by the end of grade 4 and in Music and You by the end of grade 6. Similarly, melodic characteristics 7 and 9 are introduced formally in Music and You by grades 5 and 6 respectively. Melodic characteristic 5 is not found in Exploring Orff and melodic characteristics 4 and 7 are not referenced in the Elementary Music Curriculum Guide.

Item Specifications. The first 11 of the 12 melodic characteristics identified in the educational literature (see Table 1) were relevant to the present study. Melodic

Table 1

Melodic Characteristics Introduced by the End of Grade 4

	Melodic Characteristics
1	Melodies have contour.
2	Melodic motion consists of notes which repeat or move by step or skip.
3	Melodies may be performed or read in terms of tonic solfa (do, re, mi, fa, so, la, ti).
4	Melodies may be performed or read in terms of absolute letter names (A, B, C, D, E, F, G).
5	Diatonic scales are comprised of major and minor seconds.
6	Melodies may be perceived visually in written notation.
7	Melodies may be heard when the sound is not physically present (termed inner hearing or audiation).
8	Melodies have a tonal center.
9	Melodies may be composed of sequences.
10	A note preceded by a flat is played a semi-tone lower than the corresponding unaltered tone. A note preceded by a sharp is played a semi-tone higher than the corresponding unaltered tone.
11	The process of representing aurally perceived melodies in notational form is termed melodic dictation.
12	Original melodies may be written in traditional notation.

Table 2

Melodic Characteristics Related To Sources Consulted

Melodic Characteristic	Educational Sources				
	Musi-Canada	Music and You	Kodaly Method	Exploring Orff	Alberta Curriculum
1	x	x	x	x	x
2	x	x	x	x	x
3	x	x	x	x	x
4	x	x	x	x	
5	x	x	x		x
6	x	x (6 I)	x	x (I)	x
7	x	x (5)	x	x	
8	x	x	x	x	x
9	x	x (6)	x	x	x
10	x	x	x	x	x
11	x	x	x	x	x
12	x	x	x	x	x

Note. Numbers in parentheses indicate grade levels beyond grade four. I = inferred from content.

characteristic 12 was not addressed because it refers more appropriately to creativity than to music reading proficiency. These statements were transformed into instructional objectives for which test items were constructed (see Table 3). This was done by following the practise of beginning each statement with the phrase "the student will". To this was added the behavior to be learned or acquired by the student following relevant opportunities to learn. The breadth of melodic characteristic 6 was subdivided into three objectives: first, the ability to follow a musical score; second, the ability to discern differences between an aurally-presented melody and a single visual representation of the aural stimulus; and third, the ability to differentiate an aurally presented melody from three visual representations of the aural stimulus.

Number of Items. Because the test constructed in this study was designed for administration in a single 30 minute music class, it was necessary to limit the number of items. Therefore, each instructional objective was represented by two test items. Each pair of items is referred to as a test cluster.

Melodies Composed for Item Development

Educational Sources

The melodic material included in the selected educational sources was examined in order to identify the elements of structure found in this music by the end of grade 4. Three aspects of structure were identified in this analysis: scale, meter, and rhythm. The particulars of each element are summarized below.

Scale Structure. Selected educational sources were examined in order to identify the scales in which music is presented visually to students in the upper-elementary grades. In this respect, the grade in which the scales identified in this analysis appear to be first introduced are summarized in Table 4. This examination revealed that, in elementary school music classes, students have been introduced to the 'do' centered pentatonic scales

Table 3

Melodic Characteristics Transformed To Instructional Objectives

MC	Cluster	Item	Instructional Objective The student will:
1	1	1 2	• identify the diagram in which the contour is the same as the phrase he/she hears.
2	2	3 4	• identify pairs of notes which repeat, step, or leap.
3	3	5 6	• write the solfa syllables under the notes.
4	4	7 8	• write the note names under the notes.
5	5	9 10	• identify pairs of notes that form a semi-tone.
6	6	11 12	• identify the last aurally presented note in the melody.
6	8	15 16	• choose the measure of a notated example which differs from the melody he/she hears.
6	9	17 18	• identify the notated example which is the same as the melody he/she hears.
7	7	13 14	• identify the missing notes in an aurally presented melody.
8	10	19 20	• complete a given melody so it ends on the tonal center.
9	11	21 22	• identify from among three options the phrase in which the melodic sequence is repeated.
10	12	23 24	• identify the notated example which is the same as the melody he/she hears.
11	13	25 26	• write the notational equivalent of the melody he/she hears.

Note. MC = melodic characteristic.

of 'C', 'G', and 'F'; the 'la' centered pentatonic scales of 'a', 'e', and 'd'; the major diatonic scales of 'C', 'G', and 'F'; and the minor diatonic scales of 'a', 'e', and 'd'.

Table 4

Grade at Which Scales are First Introduced in the Educational Literature

Scales	Educational Sources				
	Musi-Canada	Music and You	Kodály Context	Exploring Orff	Alberta Curriculum
C pentatonic	1	1	2	2	I
G pentatonic	3	3	2	2	I
F pentatonic	1	2	2	2	I
a pentatonic	3	2	3	3	I
e pentatonic	3	4	3	3	I
d pentatonic	3	2	3	3	I
C diatonic	3	6	4	4	4
G diatonic	4	6	4	4	5
F diatonic	4	6	4	4	5
a diatonic	4	6	5	4	6
e diatonic	4	6	5	4	5
d diatonic	4	6	5	4	5

Note. Numbers indicate the grade at which scales appear to be first introduced. Uppercase letters denote major keys; lowercase letters denote minor keys. I = inferred from content.

Metric Structure. The metric material identified in the analysis of literature is presented in Table 5. As shown in this table, students are introduced to the meters of 2/4, 3/4, 4/4, and 6/8 by the end of grade 4 in all of the sources consulted.

Rhythmic Structure. The rhythmic material identified in the analysis of the selected literature is presented in Table 6. As illustrated in this table, eight rhythmic elements were selected for use in this study. These elements were found by the end of grade 4 in all of the sources consulted.

Table 5

Grades at which Meters are First Introduced in the Educational Literature

	Educational Sources				
Meter	MusiCanada	Music and You	Kodály Context	Exploring Orff	Alberta Curriculum
2/4	3	2	1	2	2
3/4	3	2	4	3	2
4/4	3	3	2	3	3
6/8	4	4	2	4	4

Note. Numbers indicate the grade at which scales appear to be first introduced.

Table 6

Grades at which Rhythms are First Introduced in the Educational Literature

	Educational Sources				
Rhythms	MusiCanada	Music and You	Kodály Context	Exploring Orff	Alberta Curriculum
♩	1	1	1	1	1
♫	1	1	1	1	1
𝄽	1	1	1	1	1
𝅗𝅥	2	2	2	2	2
𝅗𝅥.	2	2	4	3	3
♩	3	2	3	4	4
♩. ♪	3	4	3	4	4
♪♪♪	4	4	4	4	4

Note. Numbers indicate the grade at which scales appear to be first introduced.

Melodies Composed for this Study

Music composed for this study was based upon the particulars of scale, meter, and rhythm identified in the analysis of selected educational sources. The representation of each element in the 26 items constructed for this study (13 test clusters each represented by 2 items) is summarized in Table 7. The representation of each element is summarized below.

Scale Structure. As summarized in Table 7, for each cluster, one item was composed in a do-pentatonic or a major-diatonic scale and another item was composed in a la-pentatonic or a minor-diatonic scale. There were two exceptions to this guideline. First, pentatonic melodies do not contain minor seconds. Therefore, each item which required students to identify minor seconds (cluster 5: items 9 and 10) was composed in a diatonic scale. Second, due to the difficulty of writing sequences in the pentatonic scale, items in cluster 11 (items 21 and 22) were limited to diatonic scales.

Metric Structure. As illustrated in Table 7, of the 26 melodies composed for this study, 18 were written in 4/4 meter, 4 were written in 3/4 meter, 3 were written in 6/8 meter, and 1 was written in 2/4 meter.

Rhythmic Structure. As illustrated in Table 7, of the 26 items composed for this study, 20 items employed melodies with quarter notes, 18 employed melodies with half notes, and 15 employed melodies with eighth notes. Three items employed rhythmic configurations typically found in 6/8 meter. Dotted half notes and dotted eighth notes were used in 5 and 2 melodies, respectively. Two items employed a quarter rest.

Psychological Representation

Psychological aspects of melodic perception considered in item development were reviewed in Chapter 2. In the present study, tonal clarity (Croonen, 1994; Croonen & Kopp, 1989) was achieved by adhering to the psychological (Krumhansl, 1979) and theoretical (Piston & DeVoto, 1987) descriptions of tonal function. With the exception of

Table 7

Scalar, Metric, and Rhythmic Representation of Test Items

Cluster	Item	Scale	Meter	Rhythmic Material Referenced to Meter								
				𝅗𝅥. 3/4	𝅗𝅥 2/4, 3/4, 4/4	♩. 6/8	♩ 3/4, 4/4	♩♪ 4/4	♫ 2/4, 4/4	♫♪ 6/8	♩♪ 6/8	𝄽 6/8
1	1	C	4/4		x		x					
1	2	a	4/4	x								
2	3	d	4/4		x			x				
2	4	F	4/4		x		x		x			
3	5	C	4/4		x		x					
3	6	d	4/4				x		x			
4	7	F	4/4		x		x		x			
4	8	e	4/4				x		x			
5	9	a	3/4	x	x		x					
5	10	g	3/4	x	x		x					
6	11	e	4/4		x		x		x			
6	12	G	6/8				x		x	x	x	x
7	13	F	3/4	x	x		x		x			
7	14	a	6/8			x				x		
8	15	d	4/4		x		x		x			
8	16	G	4/4		x		x		x			

Note. Uppercase letters denote major keys; lowercase letters denote minor keys.

Table 7 (continued)

Scalar, Metric, and Rhythmic Representation of Test Items

Cluster	Item	Scale	Meter	Rhythmic Material Referenced to Meter								
				𝅗𝅥. 3/4	𝅗𝅥 2/4, 3/4, 4/4	♩. 6/8	♩ 3/4, 4/4	♩. ♪ 4/4	♫ 2/4, 4/4	♪♪♪ 6/8	♩ ♪ 6/8	𝄽 6/8
9	17	C	4/4		x		x		x			
9	18	e	6/8			x				x		x
10	19	a	4/4		x		x		x			
10	20	C	3/4	x	x		x	x				
11	21	G	2/4		x				x			
11	22	d	4/4		x				x			
12	23	e	4/4		x		x		x			
12	24	C	4/4		x		x		x			
13	25	d	4/4				x					
13	26	G	4/4				x					

Note. Uppercase letters denote major keys; lowercase letters denote minor keys.

test items which assessed the ability to identify flats and sharps, the only accidental was the raised seventh scale degree of the harmonic minor scale. Nondiatonic tones were deemed unsuitable because they would be beyond the musical background of the majority of students whose only experience with music reading was gained in school music classes. In addition, no atonal melodies were employed in the current study. In consideration of Boltz (1989a, 1989b) and Cuddy, Cohen, and Mewhort (1981), melodies represented complete phrases which progressed to the tonic—thus ending with a sense of finality. Considering the conflicting opinions with regard to how the number of directional changes influence perception (Boltz & Jones, 1986; Long, 1977), melodies represented a variety of contour patterns.

Song Text

Melodies composed for the items developed in this study were presented in the absence of a song text. The decision to proceed in this manner was based upon observations recorded in the research literature. First, Gordon (1965) observed that the use of pre-existing song material offered students familiar with the music an advantage over those students who were hearing a particular song for the first time. Second, Hudson (1986), who used songs from music series books in a pilot study, observed children singing or humming along with the familiar material as it was played on the piano. As a result, subsequent versions of Hudson's (1986) test were based upon original compositions. In consideration of this, the melodies employed in the current study were composed by the test author.

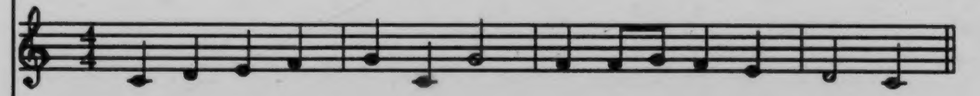
Item Format

Two item formats were used in item construction—multiple-choice and written response. For items in the multiple-choice format, students were required to choose the correct response from among the given options. For items in the written response format, students were required to write their responses as indicated, either on the given music notation or on the blanks provided below the notes. Examples of these two item formats are given in Figures 1 and 2. The distribution of item formats among test clusters is presented in Table 8.

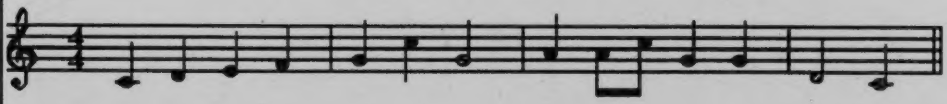
You will hear a melody sung two times. Choose the one melody which is the same as the melody you hear. Write your answer on the blank provided.

____

A.



B.



C.

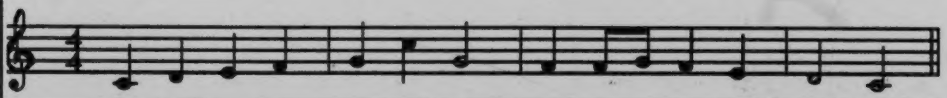


Figure 1. Item 17: Multiple-choice format.

You will hear the melody written below sung one time. Write the correct solfa syllables (d,r,m,f,s,l,t) on the lines underneath the notes.

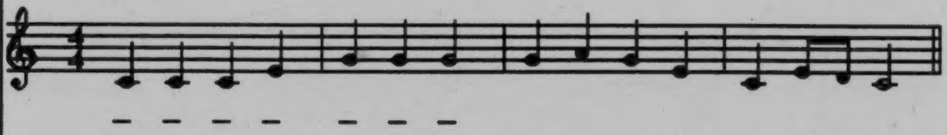


Figure 2. Item 5: Written response format.

Table 8

Categorization of Test Items: Item Format and Mode of Response

Item Format	Mode: Both Aural and Visual Necessary	Mode: Only Visual Necessary
Multiple-Choice	1,2,15,16,17,18,23,24	21,22
Written Response	11,12,13,14,19,20,25,26	3,4,5,6,7,8,9,10

Note. Numbers refer to item numbers 1 to 26.

Sequential Organization of Test Items

The order in which the initial test items were presented in the form of a test was devised in consideration of two factors—anticipated item difficulty and item format. The test began with a practise question, the structure of which was similar to items in Cluster 7. It was anticipated that students would be unfamiliar with the structure of items referenced to this objective. As such, it seemed reasonable to provide students with a practise question as a means to illustrate how to proceed with similar questions. The practise question was followed by what was believed to be the eight easiest items. The two multiple-choice items in this group were placed first, followed by six items constructed in the written response format. To avoid changing item format too often and following item construction rules that suggest items in the same format should be grouped together, another group of six written response items came next. This was followed by four items written in the multiple-choice format. The test concluded with the eight items believed to be the most difficult. Of these, four items were constructed in the written response format and four items were in the multiple-choice format. As summarized in Table 8, a total of 16 items were constructed in the multiple-choice format and 8 items were constructed in the written response format. The initial test version is presented in Appendix A.

Aural Presentation

Mode of Response. As summarized in Table 8, the test items were grouped according to mode of response and item format. Mode of response refers to those items in which students must attend to both the aural and notational component in order to correctly solve the task as opposed to those items in which students may correctly solve the task by attending solely to the visual component. There is no research literature to indicate whether students would rely on the visual component to the exclusion of the aural stimulus. Therefore, for purposes of the preliminary investigations, all melodies were sung regardless of the task to which they were referenced. Items in which only the visual component was necessary for successful completion were heard once. Items in which students were required to attend to both the aural and visual component were heard twice. The exceptions to this were items in which students were asked to follow music notation by indicating which notes were heard or not heard. It is possible that students who heard these melodies twice might be more likely to count the notes rather than attend to aspects of the tonal and/or rhythmic structure of the melody. Therefore, these items were sung only once.

Vocal Component. The test constructed in this study was designed for administration in a single 30 minute music class. In order to account, in part, for variations among the students' ability to read written text, all instructions were spoken on an accompanying cassette tape. Pre-recorded melodic examples were sung by a soprano voice. This decision was based upon the findings of Small and McCachern (1983) and Yarbrough (1991) who observed that children match pitch more accurately in response to a female vocal model as opposed to a male vocal model. Melodies were sung at a moderate tempo (quarter note = 76 beats/minute in 4/4 meter). As illustrated in Figure 3, the vocalist sang one measure on the starting pitch to cue the beginning of each melody.

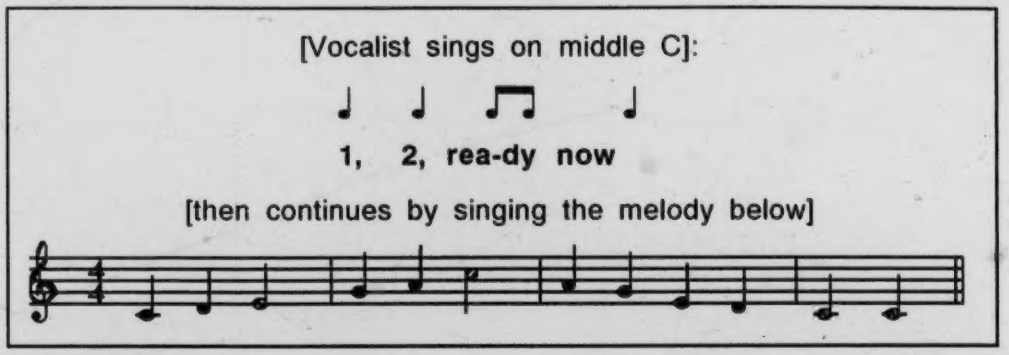

[Vocalist sings on middle C]:
1, 2, rea-dy now
[then continues by singing the melody below]

Figure 3. Aural cue.

Time Allocated for Item Completion

Total administration time of the test was 30 minutes, the typical length of time of a single music class in the schools in which the test was administered (Field Studies 1, 3, 5, and 6). This allowed for 25 minutes of actual testing time and 5 minutes for pre-administration tasks such as the distribution of tests.

In formulating the cassette tape, a period of silence followed the item directions and accompanying melody in order to allow students time to complete each item. The time allowed for each item was based on the test constructor's judgement regarding the nature of each task. These time allotments, which ranged from 5 to 30 seconds, are summarized in Table 9. Between 5 and 30 seconds of silent item completion time were allowed for items in which students supplied a written response. In the case of items 5 to 8, it was anticipated that students might need more time to write absolute note names or solfa syllables under the notes given the length of the melodies. Therefore, as stipulated in the item directions, students provided this information for notes in the first half of the melody. It was anticipated that students would complete items 11 to 13 as they listened to the melody, thereby negating the demand for silent completion time.

As with items requiring a written response, the amount of time necessary to complete multiple-choice items related to the demands of the task. First, items 1, 2, 15,

Table 9

Period of Silence Allocated for Item Completion

Item Numbers	Item Format	Task	Repeated	Silent Time (seconds)
1	MC	contour	yes	5
2	MC	contour	yes	5
3	WR	repeat	no	15
4	WR	step	no	15
5	WR	solfa names	no	30
6	WR	solfa names	no	30
7	WR	note names	no	30
8	WR	note names	no	30
9	WR	half steps	no	30
10	WR	half steps	no	30
11	WR	follow notes	no	5
12	WR	follow notes	no	5
13	WR	follow notes	no	5
14	WR	follow notes	no	5
15	MC	read notes	yes	5
16	MC	read notes	yes	5
17	MC	read notes	yes	15
18	MC	read notes	yes	15
19	WR	home tone	no	10
20	WR	home tone	no	10
21	MC	sequence	no	5
22	MC	sequence	no	5
23	MC	sharps	yes	5
24	MC	flats	yes	5
25	WR	dictation	yes	15
26	WR	dictation	yes	15

Note. MC = multiple-choice; WR = written response.

16, 17, and 18 all required students to compare a notational example to an aural representation. In items 1, 2, 15, and 16, 5 seconds separated the original presentation from its repetition. This was increased to 15 seconds for items 17 and 18 due to the anticipated difficulty level of these items. In comparison, the amount of silent time was reduced to 5 seconds for items 23 and 24 because students were required to attend only to the sharps and/or flats.

Time allowed between repetitions of the same melody, as employed in items 1, 2, 15, 16, 17, 18, 23, 24, 25, and 26, was calculated from the moment after which the melody was first performed. Time between items was calculated in the same manner. Thus, for item 1, 5 seconds separated the first performance of the melody from its repetition and another 5 seconds separated the repeated performance from the following item. In item 3, 15 seconds separated the sung melody from the following item. Time estimates are summarized in Table 9.

Content-Related Evidence of Validity

Questionnaire Responses From Elementary Music Specialists

Description of the Sample

As explained by Messick (1983), "in practise, content-related evidence [of validity] usually takes the form of consensual informed judgments about the representative coverage of content in a test and about its relevance to a particular behavioral domain of interest" (p. 38). Information is generally acquired from individuals who have personal experience with the construct under investigation (Crocker and Algina, 1986). In the present study, such data were obtained from elementary-level music teachers. Nineteen music educators employed in school jurisdictions throughout Western Canada were contacted by mail during February, 1994. Of this number, 11 chose to take part in this aspect of the study.

Participants (hereafter referred to as judges) were sent a copy of the test developed in this study, a cassette tape on which directions to students and all melodies had been recorded, an answer key, and a questionnaire referenced to the test items. They were requested to take the test (Ebel, 1956; Tenopyr, 1977), mark the test, and assess the test items by completing the accompanying questionnaire. All questionnaire responses were anonymous.

Judges were requested to submit information with reference to educational background, years of experience, and grades taught. This information is reported in Table 10. All judges had bachelor or masters degrees in music or music education. In addition, five of the judges had earned advanced certification in music teaching. Six of the judges had at least 10 years of teaching experience. With the exception of judge 6, who did not indicate grades taught, all respondents had teaching experience in the grades to which the test was referenced.

Structure of the Questionnaire

The structure of the questionnaire was governed, in part, by the framework of the test. The 26 test items were presented as pairs in 13 test clusters. Within each cluster, four questions were referenced to each test item:

1. Reading Level: Is the reading level of the item appropriate for students in grade 5?
2. Item Clarity: Is the item free of irrelevant material?
3. Item Relevance: How well do you think the characteristic of melody, illustrated in the sample test items, relates to melodic perception as it is portrayed in elementary school music programs?
4. Item Representativeness: This test assesses music reading skills within a group situation. With this in mind, how well do you think the sample test items assess music achievement in terms of the characteristic of melody?

In questions 1 and 2, judges assessed reading level and the inclusion of irrelevant material by indicating 'Y' for 'yes', 'N' for 'no', and '?' for 'unsure'. In questions 3 and 4, judges assessed item relevance and representativeness by means of a 5-point rating scale:

Table 10

Questionnaire Judges: Educational Background, Years of Experience, and Grades Taught

Judge	Degrees/Certificates	Years Taught	Grades Taught
1	B.Ed., A.C.	>10	1, 2, 3, 4, 5, 6
2	B.Ed., A.C.	>10	1, 4, 5, 6
3	B.Mus.	6 to 10	4, 5, 6,
4	B.Ed., B.Mus., A.C.	>10	1, 2, 3, 4, 5, 6
5	B.Mus., M.Ed., A.C.	0 to 5	1, 2, 3, 4, 5, 6
6	M.Ed.	>10	not given
7	B.Ed.	>10	3, 5, 6
8	B.Ed., B.Mus., M.Ed., A.C.	6 to 10	1, 2, 3, 4, 5, 6
9	B.Ed.	>10	1, 2, 3, 4, 5
10	B.Ed.	0 to 5	4, 5, 6
11	B.Ed., A.C.	6 to 10	4, 5, 6

Note. B.Ed = Bachelor of Education; B.Mus = Bachelor of Music; M.Ed. = Master of Education; A.C. = Advanced Certification.

1-poor, 2-fair, 3-good, 4-very good, and 5-excellent. Judge 6 created a new rating scale by using question marks in four of the items, two each for questions 3 and 4. Additionally, Judges 6 and 7 each omitted two items referenced to question 3. Considering that responses were anonymous, it was not possible to obtain further information from these judges. In all cases, responses were entered in boxes provided on the questionnaire form. An abbreviated form of the questionnaire, showing these 4 questions with reference to a single test cluster, is presented in Appendix B.

Summary of Responses

Reading Level. As summarized in Table 11, with the exception of item 9, at least 9 of 11 judges (> 80%) indicated that the reading level of items 1 to 22 was appropriate for students in grade 5. At least 7 judges indicated that the reading level was appropriate for students in grade 5 for items 23 to 26.

Item Clarity. As summarized in Table 12, at least 9 of 11 judges (> 80%) indicated that items 1 to 8, 11 to 22, and 25 and 26 were free of irrelevant material. Eight judges indicated that items 9, 10, 23, and 24 were free of irrelevant material.

Item Relevance. As summarized in Table 13, for all items at least 9 of 11 judges (> 80%) rated item relevance as good (3), very good (4), or excellent (5)—thus indicating that items were relevant to the specified content domain. The mean score for each item was equal to or greater than 4.0; the medians were 4, 4.5, or 5. Overall, then, it was concluded that each of the items was relevant to the instructional objective to which it was referenced.

Deviation scores obtained for Judges 1 and 5, 35 and 40 respectively, indicated that the responses obtained from these two judges differed from the responses obtained from the remaining nine judges. Considering that Judges 1 and 5 provided written comments relevant to item revision, their responses were included in the analysis.

Item Representativeness. As summarized in Table 14, for all items at least 9 of 11 judges (> 80%) rated item representativeness as good (3), very good (4), or excellent (5). The mean score for each item was equal to or greater than 4.0; the medians were 4, 4.5, or 5. Overall, it was concluded that the set of 26 items was representative of characteristics of music reading as portrayed in music classes at the grade 5 level.

Deviation scores obtained for Judges 1 and 5, 38 and 30 respectively, indicated that the responses obtained from these two judges differed from the responses obtained from the remaining nine judges. Considering that Judges 1 and 5 provided written comments relevant to item revision, responses from these judges were again included in the analysis.

Table 11

Questionnaire Responses: Reading Level (question 1)

		Judges											Totals		
C	I	1	2	3	4	5	6	7	8	9	10	11	Y	N	?
1	1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
2	3	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	10	1	
	4	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	10	1	
3	5	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	6	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	10	1	
4	7	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	8	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
5	9	Y	Y	Y	Y	N	?	?	Y	Y	Y	Y	8	1	2
	10	Y	Y	Y	Y	Y	?	?	Y	Y	Y	Y	9		2
6	11	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
7	13	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	14	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	10	1	
8	15	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	16	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
9	17	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	18	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
10	19	Y	Y	Y	Y	Y	?	Y	Y	Y	Y	Y	10		1
	20	Y	Y	Y	Y	Y	?	Y	Y	Y	Y	Y	10		1
11	21	N	Y	Y	Y	?	Y	Y	Y	Y	Y	Y	9	1	1
	22	N	Y	Y	Y	?	Y	Y	Y	Y	Y	Y	9	1	1
12	23	?	Y	Y	Y	N	Y	?	Y	Y	Y	Y	8	1	2
	24	?	Y	Y	Y	N	Y	?	Y	Y	Y	Y	8	1	2
13	25	?	Y	Y	Y	?	?	Y	?	Y	Y	Y	7		4
	26	?	Y	Y	Y	?	?	Y	?	Y	Y	Y	7		4

Note. Responses were made from three choices (Y = yes; N = no; ? = unsure). C = cluster number; I = item number.

Table 12

Questionnaire Responses: Item Clarity (question 2)

		Judges											Totals		
C	I	1	2	3	4	5	6	7	8	9	10	11	Y	N	?
1	1	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	10	1	
	2	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	10	1	
2	3	Y	Y	Y	Y	Y	?	Y	Y	Y	Y	Y	10		1
	4	Y	Y	Y	Y	Y	?	Y	Y	Y	Y	Y	10		1
3	5	Y	Y	Y	Y	Y	?	Y	Y	Y	Y	Y	10		1
	6	Y	Y	Y	Y	Y	?	Y	Y	Y	Y	Y	10		1
4	7	Y	Y	Y	Y	Y	?	Y	Y	N	Y	Y	9	1	1
	8	Y	Y	Y	Y	Y	?	Y	Y	N	Y	Y	9	1	1
5	9	?	Y	Y	Y	Y	?	Y	Y	N	Y	Y	8	1	2
	10	?	Y	Y	Y	Y	?	Y	Y	N	Y	Y	8	1	2
6	11	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	10	1	
	12	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	10	1	
7	13	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	10	1	
	14	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	10	1	
8	15	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	16	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
9	17	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	10	1	
	18	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	10	1	
10	19	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	20	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
11	21	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
	22	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	11		
12	23	Y	Y	Y	Y	-	Y	N	?	Y	Y	Y	8	1	1
	24	Y	Y	Y	Y	-	Y	N	?	Y	Y	Y	8	1	1
13	25	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	10		
	26	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	10		

Note. Responses were made from three choices (Y = yes; N = no; ? = unsure). C = cluster number; I = item number; - = omit.

Table 13

Questionnaire Responses: Item Relevance (question 3)

		Judges											Summary Statistics		
C	I	1	2	3	4	5	6	7	8	9	10	11	M	Md	R
1	1	3	5	4	4	4	5	5	5	4	5	4	4.4	4	3
	2	3	5	4	4	4	5	5	5	4	5	4	4.4	4	3
2	3	3	5	4	5	4	5	3	5	4	4	5	4.3	4	3
	4	3	5	4	5	4	5	3	5	4	4	5	4.3	4	3
3	5	3	5	5	5	2	5	4	5	3	5	4	4.2	5	4
	6	3	5	5	5	2	5	2	5	3	5	4	4.0	5	4
4	7	3	5	4	5	4	5	4	5	3	5	5	4.4	5	3
	8	3	5	4	5	5	5	4	5	3	5	5	4.5	5	3
5	9	3	5	5	5	2	5	-	5	5	5	4	4.4	5	4
	10	3	5	5	5	4	5	-	5	5	5	4	4.6	5	3
6	11	4	5	4	5	4	?	5	5	1	5	4	4.2	4.5	5
	12	4	5	4	5	4	?	5	5	1	5	4	4.2	4.5	5
7	13	4	5	4	4	3	4	5	5	1	5	4	4.0	4	5
	14	4	5	4	4	5	4	5	5	1	5	4	4.2	4	5
8	15	3	5	4	5	5	5	5	5	5	4	5	4.6	5	3
	16	3	5	4	5	5	5	5	5	5	4	5	4.6	5	3
9	17	4	3	4	5	4	5	5	5	5	5	4	4.5	5	3
	18	4	3	4	5	3	5	5	5	5	5	4	4.4	5	3
10	19	4	4	5	5	5	5	5	5	5	4	4	4.6	5	2
	20	4	4	5	5	5	5	5	5	5	4	4	4.6	5	2
11	21	3	5	4	5	2	5	5	5	5	5	5	4.5	5	4
	22	3	5	4	5	2	5	5	5	5	5	5	4.5	5	4
12	23	3	5	5	5	1	-	5	5	5	5	4	4.3	5	5
	24	3	5	5	5	1	-	5	5	5	5	4	4.3	5	5
13	25	4	5	4	5	1	2	5	5	5	5	4	4.1	5	5
	26	4	5	4	5	1	3	5	5	5	5	4	4.2	5	5
D		35	13	11	3	40	9	15	7	21	9	15			

Note. Judgements were made on a 5-point scale (1 = poor; 2 = fair; 3 = good; 4 = very good; 5 = excellent). ? = unsure; - = omit; C = cluster number; I = item number; M = median; Md = median; R = range; D = deviation.

Table 14

Questionnaire Responses: Item Representativeness (question 4)

		Judges											Summary Statistics		
C	I	1	2	3	4	5	6	7	8	9	10	11	M	Md	R
1	1	3	5	4	4	5	4	5	5	4	5	4	4.4	4	3
	2	3	5	4	4	4	4	5	5	4	5	4	4.3	4	3
2	3	3	5	4	5	4	5	3	4	4	4	5	4.2	4	3
	4	3	5	4	5	4	5	3	4	4	4	5	4.2	4	3
3	5	2	4	5	5	4	5	4	5	5	5	4	4.4	5	4
	6	2	4	5	5	4	5	2	5	5	5	4	4.2	5	4
4	7	2	5	4	5	4	5	4	3	5	5	5	4.3	5	4
	8	2	5	4	5	5	5	4	3	5	5	5	4.4	5	4
5	9	3	5	5	5	2	5	5	5	3	5	4	4.3	5	4
	10	3	5	5	5	4	5	5	5	3	5	4	4.5	5	3
6	11	4	5	4	5	4	2	5	5	1	5	4	4.0	4	5
	12	4	5	4	5	4	2	5	5	1	5	4	4.0	4	5
7	13	4	5	5	5	4	4	5	5	1	5	4	4.3	5	5
	14	4	5	5	5	3	4	5	5	1	5	4	4.2	5	5
8	15	4	5	4	5	5	5	5	5	5	4	5	4.7	5	2
	16	4	5	4	5	5	5	5	5	5	4	5	4.7	5	2
9	17	4	3	5	5	4	5	5	5	5	5	4	4.5	5	3
	18	4	3	5	5	3	5	5	5	5	5	4	4.5	5	3
10	19	4	4	5	5	5	5	5	4	5	4	4	4.5	5	2
	20	4	4	5	5	5	5	5	4	5	4	4	4.5	5	2
11	21	3	5	4	5	4	5	5	5	5	5	5	4.6	5	3
	22	3	5	4	5	4	5	5	5	5	5	5	4.6	5	3
12	23	3	5	5	5	1	?	5	5	5	5	4	4.3	5	5
	24	3	5	5	5	1	?	5	5	5	5	4	4.3	5	5
13	25	4	5	5	5	2	4	5	5	5	5	4	4.5	5	4
	26	4	5	5	5	2	4	5	5	5	5	4	4.5	5	4
D		38	14	6	4	30	10	12	10	18	8	16			

Note. Judgements were made on a 5-point scale (1 = poor; 2 = fair; 3 = good; 4 = very good; 5 = excellent). ? = unsure; C = cluster number; I = item number; M = mean; Md = median; R = range; D = deviation.

Cluster-By-Cluster Analysis

Responses to questions 1, 2, 3, and 4, summarized in Tables 11, 12, 13, and 14, respectively, revealed that the initial analysis of the judges' comments yielded information useful to further refinement of the items. These suggestions are documented below in terms of the item clusters. Where applicable, the discussion concludes with a description of item revisions for each cluster.

Cluster 1 (items 1 and 2)

The test items referenced to cluster 1 were designed to assess a student's ability to identify the contour of a given melody.

Questionnaire Responses. Eleven and 10 judges, respectively, responded affirmatively to questions 1 and 2. Ratings for questions 3 and 4 ranged from 3 to 5 with a median of 4 indicating that the judges viewed these items as relevant to and representative of melodic perception as portrayed in elementary school music programs.

Suggestions For Item Revision. For item 2, Judges 4 and 5 observed that the W-shape presented in written form did not correspond exactly to the aural version. As shown in Figure 4, Judges 4 and 5 suggested the melody presented aurally corresponded to a reduced or "lop-sided" W (frame 2) in contrast to that suggested (frame 1). To correct this, the melody presented at the top of the figure was changed to the melody shown at the bottom of the figure.

Cluster 2 (items 3 and 4)

The test items referenced to cluster 2 were designed to assess a student's ability to identify notes that repeat and notes that are separated by a step, respectively.

Questionnaire Responses. Questions 1 and 2 were each answered affirmatively by 10 judges. Ratings for questions 3 and 4 ranged from 3 to 5 with a median of 4 indicating that the judges viewed these items as relevant to and representative of melodic perception as portrayed in elementary school music programs.

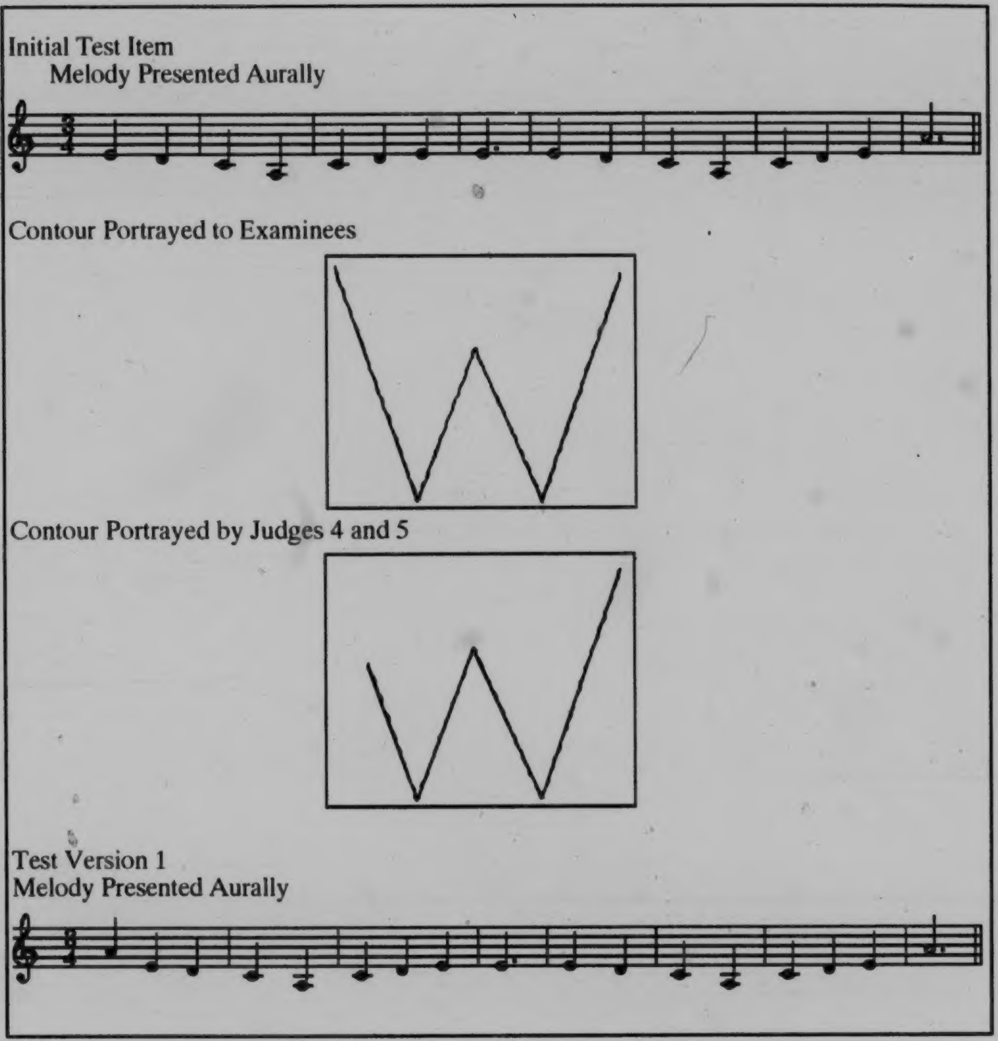

Initial Test Item
Melody Presented Aurally
Contour Portrayed to Examinees
Contour Portrayed by Judges 4 and 5
Test Version 1
Melody Presented Aurally

Figure 4. W-shaped melody.

Suggestions For Item Revision. Judge 6 indicated that the item contained irrelevant material. She thought that students might be confused by the method of presentation in which a measure is counted on the starting pitch before the melody is sung. Since no viable alternative was offered, this method of introduction was maintained for Test Version 1.

Judges 1 and 6 observed that these questions could be answered correctly by attending only to the written notation. Due to the absence of previous research regarding how students respond to this type of item, both modes of stimulus (aural, written notation) were retained at this point in the study. Judge 8 observed that the inclusion of circled notes, as illustrated in Figure 5, might be misleading. It was suggested that these notes be placed on a staff. However, it was anticipated that such an addition might supply students with the correct response. Instead, these illustrations were deleted. Judge 3 noted that students were required to circle the correct group of notes, even though directions did not indicate so. Directions to students were revised to account for this observation.

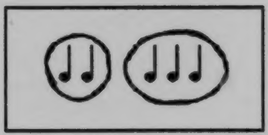


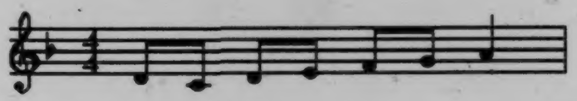
Figure 5. Circled notes as illustrated in directions to students.

Cluster 3 (items 5 and 6)

The test items referenced to cluster 3 were designed to assess a student's ability to equate notes with their corresponding solfa syllable names (do, re, mi, fa, so, la, and ti).

Questionnaire Responses. Questions 1 and 2 were each answered affirmatively by 10 judges. Ratings for questions 3 and 4 ranged from 2 to 5 with medians of 5, indicating that these items were both relevant to and representative of the specified content domain. Judge 6 did not provide a rationale for the ratings of 2 assigned to these two items.

Suggestions For Item Revision. Suggestions for item improvement offered by Judges 1 and 7 provided an explanation to the low ratings supplied by these judges. Judge 1 observed that these questions could be answered correctly by attending only to the written notation. While students might answer this item by attending to the key signature and position of the notes on the staff, at present there is no research to indicate if aural skills would be used to arrive at a solution. Hence, pending additional information, no change was made. Judge 7 thought that the items would be improved if, for each item, the syllable name of the first note was given to the students. Although it is possible that this would serve as a means by which to explain the task, at this point in the study the initial syllable was not provided because, as illustrated in Figure 6, supplying this information might negate any assumed effects of the aural stimulus.



la	so	la	ti	do	re	mi	visual response key of d minor
re	do	re	mi	fa	so	la	aural response key of C major

Figure 6. Item 6: Visual response versus aural response.

Cluster 4 (items 7 and 8)

The test items referenced to cluster 4 were designed to assess a student's ability to identify notes by their absolute letter names (A, B, C, D, E, F, and G).

Questionnaire Responses. Questions 1 and 2 were answered affirmatively by 11 and 9 judges, respectively. Ratings for questions 3 and 4 ranged from 2 to 5 and 3 to 5, respectively. The large range for item 3 was due solely to a rating of 2 assigned by Judge 1. Medians of 5 indicate that, for the most part, these items were relevant and representative.

Suggestions For Item Revision. Judges 6 and 9 both thought that the items were not free of irrelevant material because more notes were given than those for which students were required to supply the note name. Judges 1 and 8 noted that these items could be completed visually without attending to the aural stimulus. At this point in the study, it was not known if students would attend only to the visual presentation. Therefore, items 7 and 8 remained in their present form.

Cluster 5 (items 9 and 10)

The test items in cluster 5 were designed to assess a student's ability to identify places in which the notes were separated by the interval of a half-step.

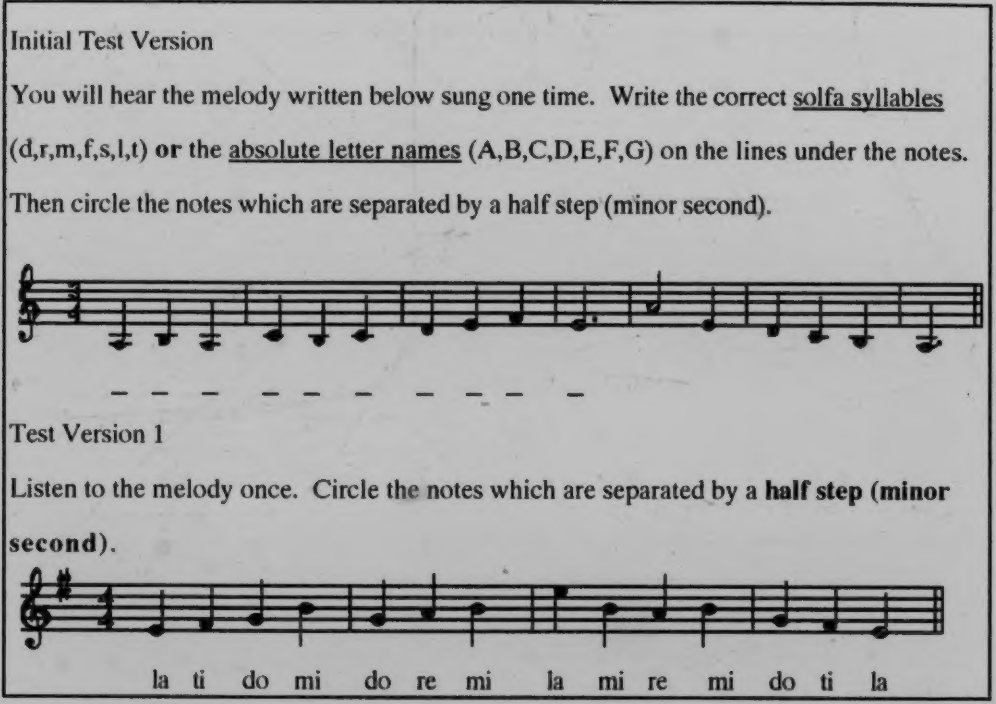
Questionnaire Responses. Questions 1 and 2 were answered affirmatively by 8 and 9 judges, respectively. Ratings for questions 3 and 4 ranged from 2 to 5 with medians of 5 indicating that these items were both relevant and representative. Judge 5, who assigned a rating of 2 to item 9 in response to question 3, and Judge 7, who omitted question 3, did not provide a reason for their responses.

Suggestions For Item Revision. Judges 1 and 8 observed that students were expected to accomplish too much in the time provided. Therefore, as illustrated in the bottom frame of Figure 7, the solfa syllables were provided in the revised form of this item. Judge 8 also noted that item 9 was made more difficult by the inclusion of ledger line notes. To address this concern, the melody employed in this item, composed in the key of 'a' minor, was changed to a melody composed in the key of 'e' minor (see Figure 7).

Cluster 6 (items 11 and 12)

The test items referenced to cluster 6 were designed to assess a student's ability to follow music notation. In these questions, only part of the melody was sung aloud. Students were then required to circle the final note they heard.

Questionnaire Responses. Eleven and 10 judges, respectively, responded affirmatively to questions 1 and 2. Ratings for questions 3 and 4 ranged from 1 to 5 with medians of 4.5 and 4, respectively. Ratings of 1 were supplied by Judge 9. Judge 6 used question marks to rate these items in response to question 3.



Initial Test Version

You will hear the melody written below sung one time. Write the correct solfa syllables (d,r,m,f,s,l,t) **or** the absolute letter names (A,B,C,D,E,F,G) on the lines under the notes. Then circle the notes which are separated by a half step (minor second).

Test Version 1

Listen to the melody once. Circle the notes which are separated by a **half step (minor second).**

Figure 7. Item 9: Comparison of Initial Test Version and Test Version 1.

Suggestions For Item Revision. Judge 9 believed that these items assessed awareness of rhythm rather than perception of melody. Similarly, Judge 6 observed that students could answer these items correctly by counting the number of notes which were sung. It was anticipated that the rhythmic structure would influence how students solved these items. However, at this point in the study, it was not possible to determine if students would attend exclusively to the rhythm. Judge 8 thought that students would complete the task successfully if they heard the melody twice. This possibility had been considered. However, it was decided that students who heard the melody twice might be more likely to count the notes rather than attend to the tonal and/or rhythmic structure of the melody. Therefore, items 11 and 12 were not revised.

Cluster 7 (items 13 and 14)

The test items referenced to cluster 7 were designed to assess a student's ability to follow music notation. In these questions the aural stimulus was divided into three sections: the first and last were sung aloud and the second was silent. Students were required to follow the notation and indicate which notes were silent.

Questionnaire Responses. Questions 1 and 2 were answered affirmatively by 11 and 10 judges, respectively. Ratings for questions 3 and 4 ranged from 1 to 5 with medians of 4 and 5, respectively. Ratings of 1 were supplied by Judge 9.

Suggestions For Item Revision. Judge 9 believed that these items assessed awareness of rhythm rather than perception of melody. Judges 5 and 7 suggested that these items be heard twice. It was possible, though, that this would enable students to solve these tasks by counting the number of notes, a concern expressed by Judge 6. Judge 6 also suggested the use of familiar melodies. However, it was felt that the use of familiar melodies may lead to problems within a testing situation (Gordon, 1965; Hudson, 1986). Therefore, the melodies composed specifically for this study were maintained.

Cluster 8 (items 15 and 16)

The test items referenced to cluster 8 were designed to assess a student's ability to identify the one measure of a given melody in which the aural presentation differed from the notational representation.

Questionnaire Responses. All judges responded affirmatively to questions 1 and 2. Responses to questions 3 and 4 ranged from 3 to 5 and 4 to 5, respectively, with median responses of 5. This suggests that these items were both relevant to and representative of melodic perception as portrayed in elementary school music programs.

Suggestions For Item Revision. No suggestions were offered for item revision.

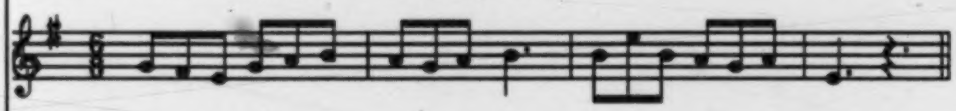
Cluster 9 (items 17 and 18)

The test items referenced to cluster 9 were designed to assess a student's ability to identify the notational equivalent of an aural melody when presented with three options.

Questionnaire Responses. Questions 1 and 2 were answered affirmatively by 11 and 10 judges, respectively. Responses to questions 3 and 4 ranged from 3 to 5 with medians of 5 indicating that these items were relevant to and representative of the content domain.

Suggestions For Item Revision. Judges 1 and 5 suggested that item 18 may be too difficult due, in part, to the number of notes employed in the melody. Therefore, as illustrated in Figure 8, the melody composed in 6/8 meter (shown in the top frame) was replaced by a composition in 4/4 meter (shown in the bottom frame).

Initial Test Version: Option A- Correct Response



Test Version 1: Option C - Correct Response

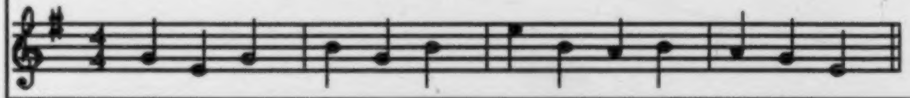


Figure 8. Item 18: Comparison of Initial Test Version and Test Version 1.

Cluster 10 (items 19 and 20)

Test items referenced to cluster 10 were designed to assess a student's ability to determine the placement of the tonal center, commonly designated the home tone of a given melody.

Questionnaire Responses. Ten and 11 judges, respectively, responded to questions 1 and 2 in the affirmative. Ratings to questions 3 and 4 ranged from 4 to 5 with a median of 5 indicating that these items were relevant and representative.

Suggestions For Item Revision. Judges 7 and 8 suggested that these items be sung twice. This recommendation was incorporated into the revised version of the test.

Cluster 11 (items 21 and 22)

The test items referenced to cluster 11 were designed to assess a student's ability to identify a repeating melodic sequence.

Questionnaire Responses. Nine and 11 judges responded affirmatively to questions 1 and 2. Ratings on questions 3 and 4 ranged from 2 to 5 and 3 to 5, with 5 being the median response. Judge 5 did not provide a rationale for ratings of 2 that she gave.

Suggestions For Item Revision. Judge 8 suggested that students should have the opportunity to complete a demonstration question before completing the task. Given the available time for test administration (see page 56), this recommendation was not incorporated into the revised test version. This suggestion was noted for future reference.

Cluster 12

The test items referenced to cluster 12 were designed to assess a student's ability to identify the sharps or flats in an aural presentation of a melody.

Questionnaire Responses. Questions 1 and 2 were answered affirmatively by 7 and 8 judges, respectively. Ratings for questions 3 and 4 ranged from 1 to 5 with medians of 5. Judge 6, who did not rate these items in response to question 3 and who provided question marks for both items in response to item 4, commented on the difficulty of these items but did not offer suggestions for improvement.

Suggestions For Item Revision. As presented in the initial test version, students heard two versions of a melody and chose the notation which was the same as the second melody they heard. Judges 1, 5, 6, and 7 thought that this format was very confusing. Judges 1 and 7 proposed that students hear one melody and choose the option which corresponded to what they heard. The test items were revised accordingly.

Cluster 13 (items 25 and 26)

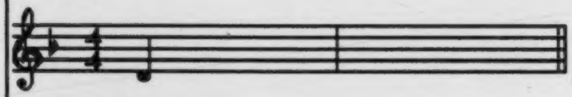
The test items referenced to cluster 13 were designed to assess a student's ability to notate an aurally dictated melody.

Questionnaire Responses. Seven judges responded affirmatively to question 1. Four judges were unsure. Ten judges responded affirmatively to question 2. Ratings for questions 3 and 4 ranged from 2 to 5 and 1 to 5 respectively, with medians of 5. The rationale supplied by Judge 5 for the low ratings given to questions 3 and 4 is discussed under suggestions for item revision.

Suggestions For Item Revision. Judges 1 and 8 suggested that students be given the first note or tonal center. In addition, Judge 5 suggested that the rhythm be notated above the staff. This is illustrated for item 25 (see Figure 9). Item 26 was revised similarly.

Initial Test Item

You will hear a melody sung two times. Write the melody on the staff below. The first note is given.



Test Version 1

Listen to the melody three times. Write the melody on the staff below. The first note and the rhythm are given.

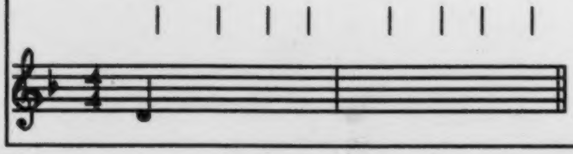


Figure 9. Item 25: Comparison of Initial Test Version and Test Version 1.

Time Allocated for Task Completion

Trial Administration

All melodies and instructions to students were recorded on a cassette tape which accompanied the test items. There was no research to guide decisions with regard to time needed to complete tasks. Therefore, in order to assess the appropriateness of time allotted, the test was administered to 26 grade 6 students attending a school in which a music specialist was responsible for music instruction. During this administration, students were instructed to raise their hands after they had completed each item. This was recorded on a checklist by both the researcher and the music specialist.

The goal of this administration was to set time limits which would allow at least 85% of the students to complete the items while staying within the total administration time of 30 minutes.

Data Analysis. The number of students who did not complete each item within the time allowed for that item is summarized in Table 15 for each item. Within this table, time allowed to complete each item for the initial test administration is compared to time allowed to complete each item in Test Version 1 (see Chapter 4). As well, in Table 15, the number of times the melodies are heard in the initial test administration is compared to the number of times that melodies are heard in Test Version 1

At least 22 out of 26 (85%) of the students completed items 1, 2, 5, 6, 9, 10, 11, 12, 13, and 14 in the initially allowed time. Time allowed to complete these items remained unchanged. For items 3, 4, 7, 8, and 23, the number of students who were unable to complete a particular item ranged from 6 to 8. The time allowed for the completion of these items was extended by 5 seconds. In order to maintain equal time requirements for similar items within clusters, the number of seconds permitted to complete item 24 was also extended. Fourteen and 12 students did not have sufficient time to complete items 19 and 20, respectively. The number of times that these items were heard was increased from once to twice with 5 seconds separating the original performance from its repetition and

Table 15

Time Allocated and Times Heard: Trial Administration and Test Version 1

			Administration			
			Trial		Test Version 1	
Item	Task	[a]Not Complete	Time (seconds)	[b]Times Heard	Time (seconds)	Times Heard
1	contour	0	5	2	5	2
2	contour	0	5	2	5	2
3	repeat	7	15	1	20	1
4	step	6	15	1	20	1
5	solfa syllables	4	30	1	30	1
6	solfa syllables	3	30	1	30	1
7	note names	7	30	1	35	1
8	note names	8	30	1	35	1
9	half step	0	30	1	30	1
10	half step	0	30	1	30	1
11	follow notation	4	5	1	5	1
12	follow notation	4	5	1	5	1
13	follow notation	2	5	1	5	1
14	follow notation	5	5	1	5	1
15	read notation	0	5	2	3	2
16	read notation	0	5	2	3	2
17	read notation	2	15	2	10	2
18	read notation	1	15	2	10	2
19	home tone	14	10	1	5	2
20	home tone	12	10	1	5	2
21	sequence	1	5	1	2	1
22	sequence	1	5	1	2	1
23	sharps	7	5	2	10	2
24	flats	3	5	2	10	2
25	dictation	8	15	2	10	3
26	dictation	3	15	2	10	3

[a] Numbers of students out of 26 who did not have time to complete the item. [b] Number of times students hear the melody.

another 5 seconds separating the repeated performance from the question that followed. Eight and 3 students indicated that time was insufficient to complete items 25 and 26, respectively. The number of times that these items were heard was increased from two times to three times with 10 seconds separating the original performance from each of its repetitions. In an attempt to maintain a total administration time of 30 minutes, total time permitted to complete items 15 and 16 was reduced by 4 seconds; total time to complete items 17 and 18 was reduced by 10 seconds.

Chapter Summary

The procedures employed in the development of the initial test items are described in Chapter 3. This process involved the review and analysis of selected educational resources which view music literacy as a fundamental goal of instruction and a review of representative psychological studies which examine the cognitive representation of melodic perception. Together, these analyses provided the rationale for the development of 26 individual test items, clustered in pairs of two, with each cluster referenced to one of the 13 instructional objectives identified in a review of selected educational materials. Content-related evidence of validity was examined by means of questionnaires forwarded to 11 subject-area specialists who were requested to take the test and then rate and, if they wished, comment on the relevance and representativeness of test items as related to the above-specified content domain. The judges were also asked to rate the reading level and clarity of the items developed in this study. In order to determine the appropriateness of the amount of time allocated to complete each item, the test was administered to 26 grade 6 students who were requested to raise their hands after they had completed each item. The purpose for this administration was to set limits that would allow at least 85% of examinees the time needed to complete each task. Revisions based on the reviews of subject-area specialists and the findings of the group administration culminated in the formulation of Test Version 1.

CHAPTER 4

FIELD STUDY 1: TEST VERSION 1 (GROUP ADMINISTRATION)

Overview

As illustrated in Chapter 3, construction of the prototype test items was based on an analysis of selected educational materials representative of course content in upper-elementary music classes and the identification of goals of instruction relative to specific aspects of music reading. Content-related evidence of validity was secured from subject-area specialists who, in response to a questionnaire, indicated that items were relevant to and representative of the specified content. Refinement of the prototype items was guided by suggestions for item improvement supplied by these judges. This was supplemented by information regarding time required to complete the tasks obtained through group administration to a class of grade 6 students. At this juncture, it was necessary to examine how students responded to the items developed in this study. This problem was addressed in Field Study 1.

Description of the Sample

Overview of Field Studies 1 Through 6

All students who participated in Field Studies 1 through 6 attended schools in which music classes were taught by subject-area specialists. The schools were located within three school districts, two public and one separate,[1] located in Central Alberta, Canada. The suburban schools were located in a metropolitan area with a population of approximately 850,000. None of the schools were located in inner-city areas.

The test items developed in this study were administered to fifth-grade students towards the end of the school-year in Field Studies 1, 2, 4, 5, and 6 and to sixth-grade students at the beginning of the school-year in Field Study 3. The test items were administered to groups of students in Field Studies 1, 3, 5, and 6 and to individuals in Field Studies 2 and 4. Participation is summarized in Table 16.

---

[1] All three school districts were government-funded. School Districts 1 and 3 (public) had no religious affiliation; School District 2 (separate) was affiliated with the Roman Catholic Church.

Table 16

Summary of Field Studies

Field Study	Type	# of Students	# of Classes	# of Schools	School District	Time Period
1	group	49	2	1	1	04/94
2	individual	8	2	2	1	05/94
3	group	77	4	4	1, 2	10/94
4	individual	10	2	2	3	02/95
5	group	23	2	2	3	03/95
6	group	214	12	7	1, 2, 3	04 - 05/95

All students participating in these studies attended regular fifth- and sixth-grade classes. Each student involved in Field Studies 1 through 6 returned a form signed by a parent or guardian indicating that the student was allowed to take part in this study (see Appendix C). Special-needs students (students with special learning considerations) did not take part in the study. An exception was a visually-impaired student who was given a copy of the test enlarged from 8 1/2 by 11 sized paper to 11 by 17 sized paper (Field Study 1).

Field Study 1

In January, 1994, the test author was invited to present the underlying premise of this study to a group of music specialists who taught in schools under the jurisdiction of School District 1. Teachers who indicated that music literacy was a fundamental goal of instruction in their music programs and who were willing to participate in this study were requested to take part in Field Study 1. One of these teachers allowed Test Version 1 to be administered in a school for which she was responsible for the music program.

Test Version 1, consisting of 26 items, was administered to students in two grade 5 classes in April, 1994. Forty-nine students completed the test. A copy of Test Version 1 is provided in Appendix D.

Item Analysis

Method

Two broad classes of test interpretation were identified in the literature: norm-referenced interpretation and criterion-referenced interpretation. A norm-referenced interpretation requires that an individual's test performance is compared to the test performance of the remaining people who took the test. A criterion-referenced interpretation requires that an individual's test performance is compared to a predetermined performance standard. All item analyses conducted in the present investigations (Field Studies 1, 3, 5, and 6) were based on a norm-referenced interpretation.

Two major criteria are used to identify how items are functioning within a norm-referenced framework: item difficulty and item discrimination. With respect to the current study, the interpretation of each of these indices is described below.

Item Difficulty

Item difficulty (p-level) refers to the proportion of students who answer a given item correctly (see Figure 10). Methods for selecting test items based on their difficulty indices depend on the purposes for which the test is constructed. As noted in Anastasi (1988), "the closer the difficulty of an item approaches 1.00 or 0, the less differential information about test takers it contributes. Conversely, the closer the difficulty level approaches .50, the more differentiations the item can make" (p. 203). Consequently, in a norm-referenced context, acceptable item difficulty generally falls within a range of medium difficulty with most of the items clustering about the .50 value to yield a mean p of approximately 50% correct. Using this guideline provides the greatest opportunity for a wide dispersion in the scores.

For purposes of the current analysis, the range of medium difficulty was defined as being within the guideline of $.20 < p > .80$. Items with difficulty indices beyond this accepted guideline were examined to determine why so many students had responded incorrectly ($p < .20$) or correctly ($p > .80$). As detailed in the item-by-item analysis, where appropriate, item revisions were undertaken as a means by which to influence the magnitude of difficulty indices obtained in subsequent group administrations. In this respect, the goal was to obtain difficulty indices which approached the value of .50.

With respect to multiple-choice items, incorrect alternatives for which the p-level was below .20 were examined in order to determine why they were chosen by a small proportion of the students.

$$P = \frac{R}{T} \times 100$$

where p = proportion of students who answered an item correctly, R = the number of students who answered the item correctly, and T = the total number of students in the sample (Gronlund and Linn, 1990).

Figure 10. Formula for item difficulty.

Item Discrimination

Item discrimination refers to the extent to which students' responses to a given item are related to their scores on the measure as a whole. For purposes of the present study, item discrimination was reported using the point-biserial correlation (r_{pbis}), a Pearson product-moment correlation coefficient which describes the relationship between a dichotomous variable (the score on a single test item) and a continuous variable (the score on the aggregated items) (Crocker & Algina, 1986). For any given item, evidence of positive discrimination requires that the mean score of students who answered the item correctly be higher than the mean score of students who answered the item incorrectly. A statistical formula used for the computation of the point-biserial correlation is shown in Figure 11.

$$\rho_{pbis} = \frac{(\mu_{+} - \mu_{X})}{\sigma_{X}} \sqrt{p/q}$$

where μ_+ is the mean criterion score for those who answer the item correctly, μ_X is the mean criterion score for the entire group and σ_X is their standard deviation, p is item difficulty, and q is (1-p) (Crocker & Algina, 1986, p. 317).

Figure 11. Formula for the point-biserial index of correlation.

With respect to the current study, correlational indices of discrimination exceeding .20 indicated that the item discriminated adequately between those students who answered the item correctly and those who responded incorrectly. Items which showed negative discrimination ($r_{pbis} < .00$) were examined closely for evidence of defects in item construction. Similarly, items for which the point-biserial correlation for the correct response ranged from .00 to .20 were examined in order to determine why the item did not discriminate adequately (Ebel & Frisbie, 1991). In this respect, it was observed that a low index of discrimination did not necessarily provide evidence of faulty item construction. As noted in Gronlund and Linn (1990), "as we move away from the 50 percent level of difficulty toward easier or more difficult items, the index of discriminating power becomes smaller. Thus, items that are very easy or very difficult have low indexes of discriminating power" (p. 254). In consideration of this, item discrimination was interpreted in terms of item difficulty.

With respect to multiple-choice items, it was expected that incorrect alternatives would discriminate negatively thus indicating that the mean score of students who answered the item incorrectly was lower than the mean score of students who answered the item correctly. Items for which the point-biserial indices approached .00 ($-.20 < r_{pbis} < .00$) were examined in order to determine why they were essentially nondiscriminating.

Field Study 1

With respect to data obtained in Field Study 1, the difficulty indices and point-biserial correlations for the correct response and incorrect options are reported in Table 17. Omitted responses and responses not corresponding to one of the options of the item, clearly indicating that students did not know how to proceed with the question, were coded "other". These results for each cluster or pair of items are described below.

Item-By-Item Analysis

Items 1 and 2

Items 1 and 2 were designed to assess a student's ability to identify the contour of a given melody.

Indices of Difficulty and Discrimination. These items had difficulty levels of .63 and .65, respectively. In both items, the linear representation of the aural stimulus was essentially nondiscriminating, being chosen by no students ($r_{pbis} = .00$) and 1 student ($r_{pbis} = -.07$), respectively.

Item Revision. As illustrated in Figure 12, the diagrams in Test Version 1 represented contour as based on the graphic illustrations of M-shaped and V-shaped contours provided by Morrongiello and Roes (1990). In revised form the diagrammatic representations of pitch were converted to traditional music notation (see Figure 12, second frame for each contour). The overall contour of the options remained unchanged. In addition, the order in which the visual representation of the contours was presented was altered in item 1 so that the contours (V-shaped, M-shaped, linear) would not be presented in the same order for both items. This was done so that students would be required to read both items with equal scrutiny. Therefore, in future studies, the correct response for both items 1 and 2 was option 2.

Table 17

Field Study 1: Indices of Difficulty and Discrimination

				Alternatives				
Item	Task	Format	Statistic	A	B	C	D	Other
1	contour	MC	p	.00	.37	.63*		
			r_{pbis}	.00	-.37	.37		
2	contour	MC	p	.04	.65*	.31		
			r_{pbis}	-.07	.28	-.26		
3	repeat	WR	p	.10*	.88			.02
			r_{pbis}	.25	-.22			-.03
4	step	WR	p	.10*	.88			.02
			r_{pbis}	.05	.11			-.15
5	solfa syllables	WR	p	.22*	.72			.06
			r_{pbis}	.52	-.52			-.03
6	solfa syllables	WR	p	.22*	.70			.08
			r_{pbis}	.59	-.51			-.04
7	note names	WR	p	.25*	.70			.08
			r_{pbis}	.59	-.28			-.36
8	note names	WR	p	.25*	.69			.06
			r_{pbis}	.48	-.37			-.15
9	half step	WR	p	.04*	.94			.08
			r_{pbis}	.42	-.31			-.07
10	half step	WR	p	.00*	.92			.08
			r_{pbis}	.00	.08			-.08
11	follow notes	WR	p	.53*	.35			.12
			r_{pbis}	.44	-.25			-.31
12	follow notes	WR	p	.51*	.41			.08
			r_{pbis}	.34	-.21			-.25
13	follow notes	WR	p	.06*	.90			.04
			r_{pbis}	.33	-.20			-.10

Note. * denotes correct response. MC = multiple-choice; WR = written response.

Table 17 (continued)

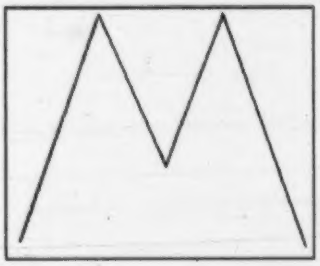
Field Study 1: Indices of Difficulty and Discrimination

				Alternatives				
Item	Task	Format	Statistic	A	B	C	D	Other
14	follow notes	WR	p	.06*	.92			.02
			r_{pbis}	.50	-.38			-.11
15	read notes	MC	p	.08	.23	.55*	.02	.23
			r_{pbis}	-.17	-.16	.24	-.07	-.03
16	read notes	MC	p	.00	.10	.63*	.08	.19
			r_{pbis}	.00	-.28	.29	-.08	-.08
17	read notes	MC	p	.47*	.35	.14		.04
			r_{pbis}	.47	-.14	-.35		-.25
18	read notes	MC	p	.43	.26	.27*		.04
			r_{pbis}	-.05	-.31	.36		-.25
19	home tone	WR	p	.45*	.49			.06
			r_{pbis}	.25	-.15			-.22
20	home tone	WR	p	.31*	.63			.06
			r_{pbis}	.42	-.32			-.17
21	sequence	MC	p	.37*	.18	.37		.08
			r_{pbis}	.15	-.04	-.04		-.15
22	sequence	MC	p	.18	.33	.47*		.02
			r_{pbis}	-.22	-.13	.35		-.19
23	sharps	MC	p	.33	.45	.18*		.04
			r_{pbis}	-.06	-.05	.27		-.25
24	flats	MC	p	.41*	.14	.43		.02
			r_{pbis}	.42	-.05	-.34		-.15
25	dictation	WR	p	.00*	1.00			.00
			r_{pbis}	.00	.00			.00
26	dictation	WR	p	.00*	1.00			.00
			r_{pbis}	.00	.00			.00

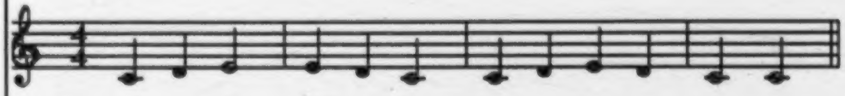
Note. * denotes correct response. MC = multiple-choice; WR = written response.

M-Shaped Contour

i. Diagram

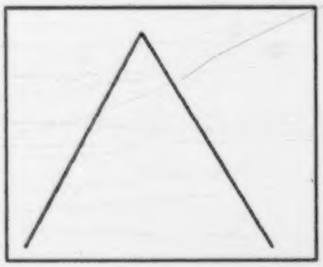


ii. Notation



V-Shaped Contour

i. Diagram



ii. Notation

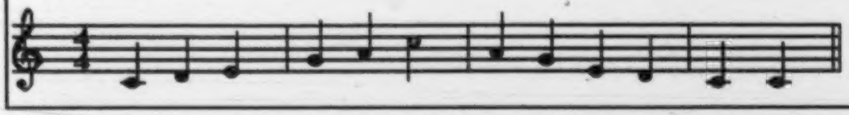


Figure 12. Item 1: Diagrammatic and notational representations of contour.

Items 3 and 4

Items 3 and 4 were designed to assess a student's ability to identify all the notes which moved by a repeat (item 3) or a step (item 4).

Indices of Difficulty and Discrimination. The difficulty of both items was .10, which is below the guideline of .20. While item 3 ($r_{pbis} = .25$) discriminated positively within accepted guidelines, item 4 ($r_{pbis} = .05$) was essentially nondiscriminating.

Examination of responses in item 3 revealed that, in general, the students' interpretation of repeated notes differed from the meaning typically assigned in the theoretical literature. As illustrated in Damschroder (1992), repeated notes, termed unisons, occur "when a pitch is struck twice in succession" (p. 3). This was the meaning apparently assumed by music educators as evidenced in the responses provided in the questionnaire. However, as revealed in the data obtained in Field Study 1, students assigned a different meaning to this term. As illustrated in Figure 13, students identified repeated notes as notes on the same line or space, regardless of the intervening interval or intervals. This was contrasted with the expected theoretical interpretation (see Figure 14).

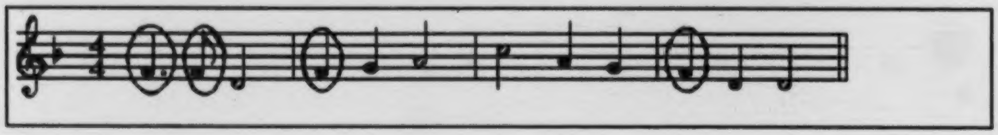


Figure 13. Repeated notes: Student interpretation.

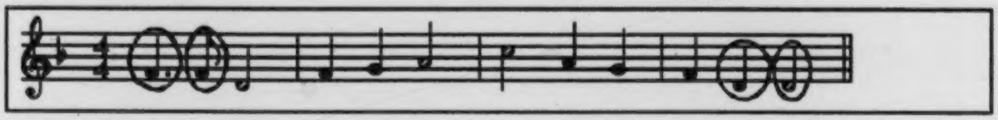
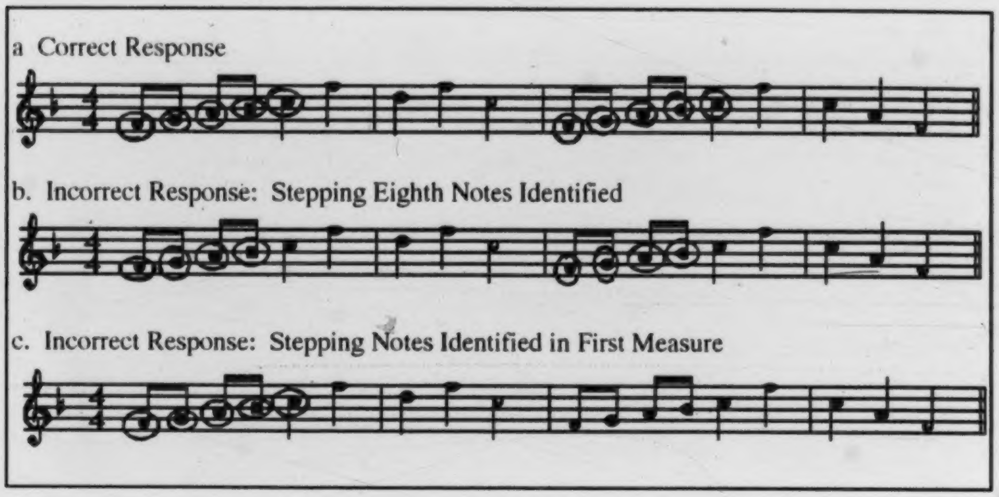


Figure 14. Repeated notes: Expected theoretical interpretation.

Item 4, which required students to identify notes separated by steps, was answered correctly by 7 students. As illustrated in Figure 15, an analysis of student responses revealed that students appeared to be influenced by the rhythmic structure of the melody. Eleven students identified the eighth notes that moved by step but failed to circle the quarter note that completed the step-wise pattern. Six students identified the notes which moved by step in the first measure but did not identify the step-wise structure in the third measure.


a Correct Response

b. Incorrect Response: Stepping Eighth Notes Identified

c. Incorrect Response: Stepping Notes Identified in First Measure

Figure 15. Item 4: Student response.

Item Revision. Data obtained for item 3 revealed a disparity between the meaning of repeated notes assigned by students and the definition of repeated notes assigned by teachers. Considering that the direct examination and analysis of pedagogical practises was beyond the scope of this study, the item in which students were requested to identify repeated notes was replaced by an item in which students were required to identify notes that move by a leap (see Figure 16). Leaps were defined as notes separated by intervals greater than a major or a minor third. This definition was used to avoid confusion with the term skips which is sometimes used to identify intervals of a major or a minor third.

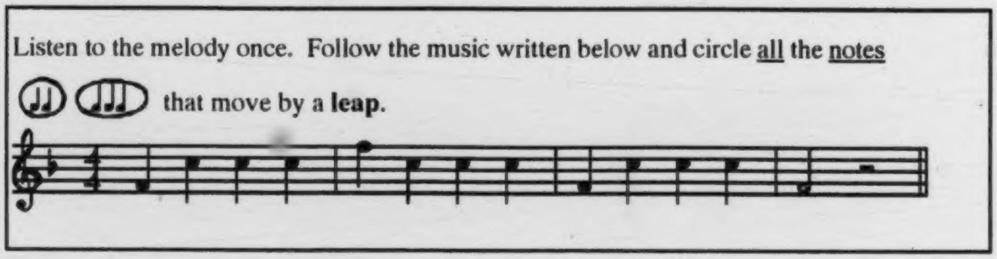

Listen to the melody once. Follow the music written below and circle all the notes that move by a **leap**.

Figure 16. Item 3: Revised form.

From a psychological perspective, responses obtained in item 4 revealed an apparent interaction between melody and rhythm. Despite its difficulty, this item remained in its present form for Field Study 2 in order to obtain more data with which to examine this phenomenon.

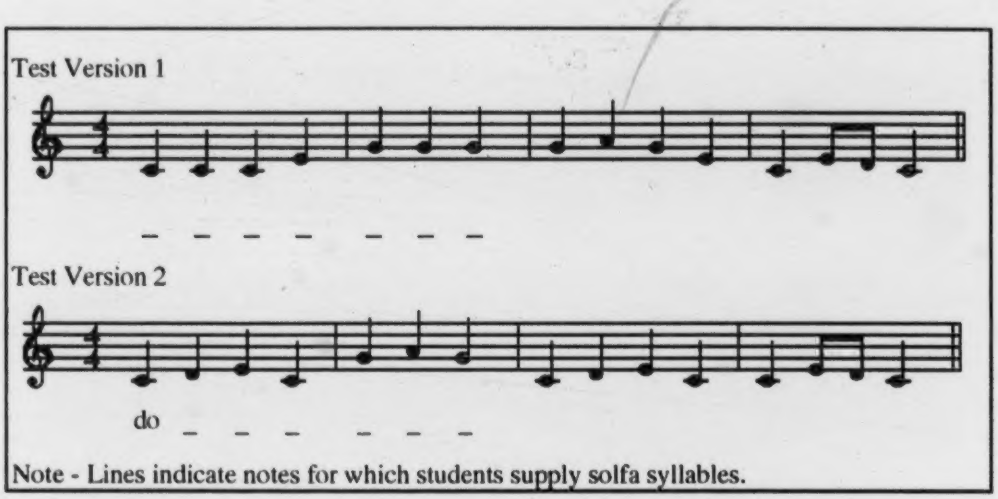
Items 5 and 6

Items 5 and 6 were designed to assess a student's ability to equate notes with their corresponding solfa syllable names (do, re, mi, fa, so, la, ti).

Indices of Difficulty and Discrimination. The difficulty index for both items, .22, was at the low end of the acceptable level. However, both items discriminated positively (r_{pbis} = .52 and .59).

Item 5 is illustrated in Figure 17. Two identifiable incorrect responses, do-do-do-re-mi-mi-mi and do-do-do-re-fa-fa-fa, were each submitted by 3 students. In each case, students misread the spaces between the notes and thus miscalculated the subsequent syllables. One student submitted the response pattern do-do-do-mi-la-la-la. The remaining 12 students began the melody on 'do'. Of these, 5 students correctly identified the repeating 'do'. The 3 students who identified the repeating 'do' as do-re-mi may have entered the syllables in the order that they appeared in the question without attending to the pattern of the notes as presented in the given melody. One student portrayed the solfa in a minor tonality. Responses from 3 students were categorized "other".

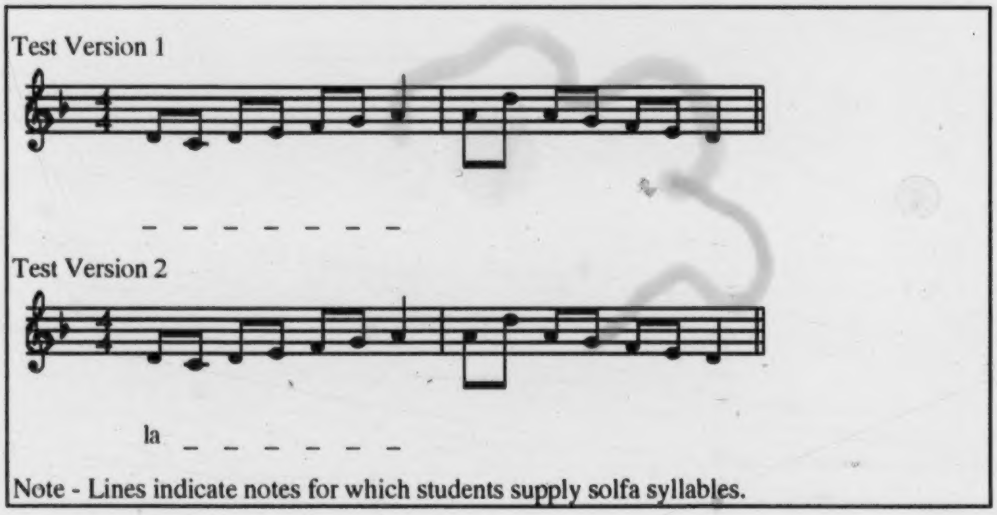
The melody employed in item 6 is illustrated in Figure 18. No students submitted the notationally correct response, la-so-la-ti-do-re-mi. Eleven students submitted the aurally correct response re-do-re-mi-fa-so-la. One student submitted a response which matched this with the exception of the final note. Thirteen students submitted responses that began on 'la', indicating a minor tonality, but none of these continued the pattern correctly. In eight of these responses 'la' was followed by 'do'. This could reflect a melodic pattern with which the students were familiar. Only 2 of these students correctly ended the melody on 'la'. As in item 5, 1 student wrote the solfa syllables in the order they appeared in the question apparently without referencing this to the order in which they appeared in the notation. Responses from 4 students were categorized "other".


Test Version 1

Test Version 2

Note - Lines indicate notes for which students supply solfa syllables.

Figure 17. Item 5: Comparison of melodies employed in Test Versions 1 and 2.


Test Version 1

Test Version 2

Note - Lines indicate notes for which students supply solfa syllables.

Figure 18. Item 6: Melodic material.

Item Revision. For both items, responses which were not described in the preceding analysis appeared to be random. Twenty and 21 students, respectively, did not assign repeated notes the same syllable name. This apparent randomness suggested that either students were unfamiliar with the task or they did not know how to proceed with the task as presented in Test Version 1. To address this problem, the solfa syllable names of the first note in the melody was supplied to illustrate the requirements of the item (see Figures 17 and 18, frame 2). The melody employed in item 5 was also revised so that students were required to identify the solfa syllable names of 4 different pitches for which the solfa syllable had not been supplied in the item. The melody for item 6 was not revised.

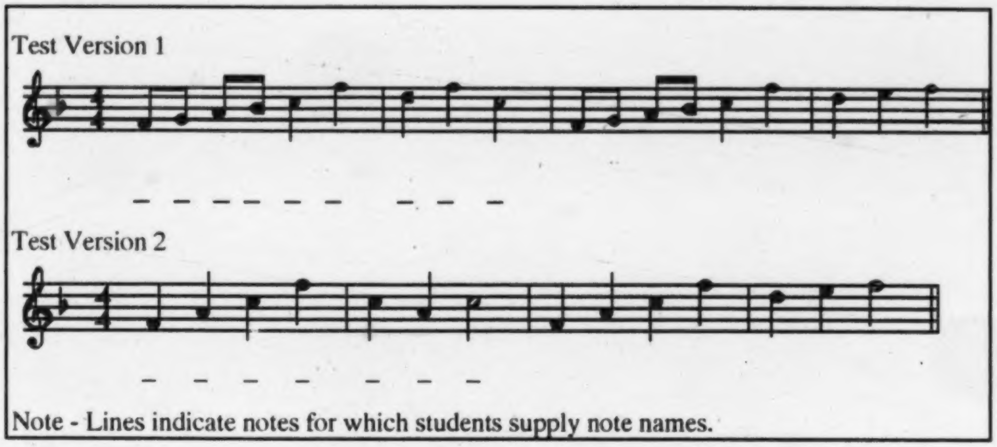
Items 7 and 8

Items 7 and 8 were designed to assess a student's ability to identify notes by their absolute letter names (A, B, C, D, E, F, and G).

Indices of Difficulty and Discrimination. Both items had difficulty indices of .25 and discriminated positively (r_{pbis} = .59 and .48, respectively).

Item 7 is illustrated in Figure 19. Five incorrect responses were identified in the data. Three students correctly identified the notes in the first measure, but named the final three notes incorrectly. One student correctly identified the notes in the spaces but omitted those on lines. Another student misread the 'F's' as 'E's' but correctly named the remaining notes. Finally, a single student consistently assigned letter names which were one tone lower than the printed note. Ten of the 18 remaining responses began on 'A' rather than 'F'. Of these, 4 responses began with the pattern A-B-C-D. Responses from 6 students could not be categorized.

Several incorrect responses were identified in item 8, the melody for which is illustrated in Figure 20. One student named the first three notes correctly. Two students placed the letter names under the notes in the order in which they appeared on the test question and 8 students began with the pattern A-B-C or A-C-D-B. Responses from 3 students could not be categorized.

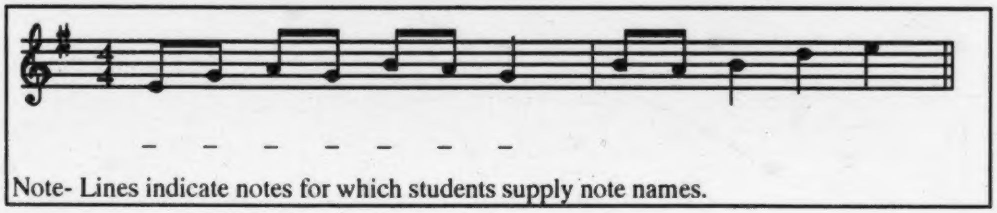


Test Version 1

Test Version 2

Note - Lines indicate notes for which students supply note names.

Figure 19. Item 7: Comparison of melodies employed in Test Versions 1 and 2.



Note- Lines indicate notes for which students supply note names.

Figure 20. Item 8: Melodic material.

The individual note names assigned by the remaining students were analyzed in terms of their frequency of occurrence as referenced to the position of the note in the melody. With respect to item 7, no distinct pattern emerged beyond the generalization that in five of the nine notes the letter name assigned most frequently was actually the next lower line or space. In item 8 the most frequent incorrect response was a line or a space below the proper name in three of nine notes. With specific reference to item 7, this pattern may be due, in part, to the structure of the melody.

A possible explanation was formulated to account for the response pattern beginning A-B-C-D. First, with respect to item 7, students may have misjudged the notes in a pattern related to the lines and the spaces of the staff. More specifically, the 'F' in the first space was read as an 'A' in the second space, the 'G' on the second line was read as a 'B' in the third space, and so on. However, this explanation was not supported by the data collected in item 8.

Item Revision. It was not known why so many responses could not be categorized (three and six responses in items 7 and 8, respectively). Therefore, other than the reduction of the number of unknown notes in item 7 from seven to four, further revisions were not made pending the information gathered from the protocol study.

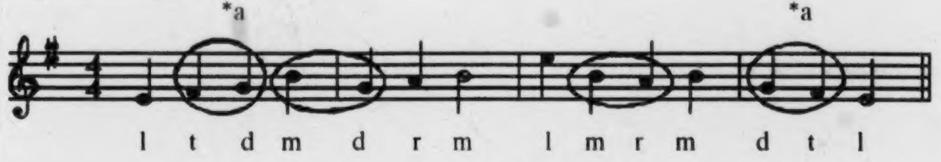
Items 9 and 10

Items 9 and 10 were designed to assess a student's ability to identify places in which the melody moved by a half-step.

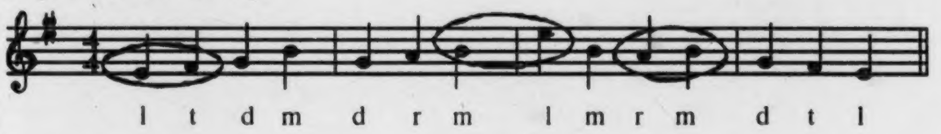
Indices of Difficulty and Discrimination. These items had p-values of .04 and .00, respectively—well below the accepted guideline of .20. One student responded correctly to item 9; no students responded correctly to item 10. For items 9 and 10, respectively, 5 and 9 students identified notes that moved by a half-step. However, these responses were incorrect because they also included notes separated by either major seconds, major thirds, and perfect fourths (with reference to item 9, see Figure 21, frame 1). Thirty-nine and 37 students, respectively, identified intervals other than a half step (with reference to item 9, see Figure 21, frame 2). Responses obtained in item 10 were similar to those obtained in item 9.

Item Revision. Revisions were not made pending the information gathered from Field Study 2.

Response includes half step



Response omits half step

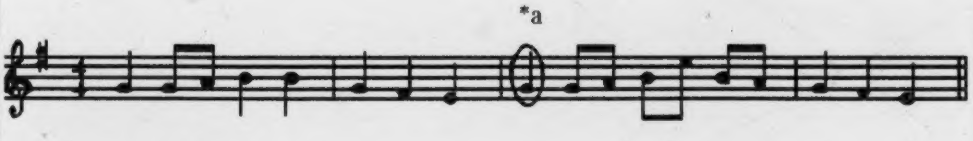


^a^Indicates half step.

Figure 21. Item 9: Student response.

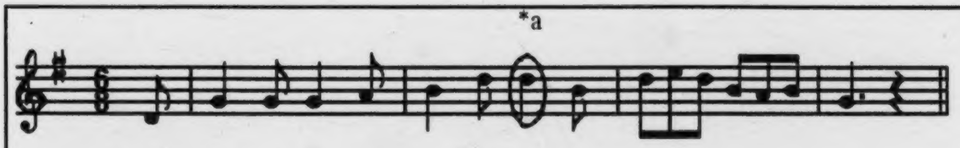
Items 11 and 12

Items 11 and 12 were designed to assess a student's ability to follow music notation. In these questions, only part of the melody was sung aloud. As illustrated in Figures 22 and 23, students were required to circle the final note they heard—referred to as the target note.



^a^Indicates target note.

Figure 22. Item 11: Target note.



^a^Indicates target note.

Figure 23. Item 12: Target note.

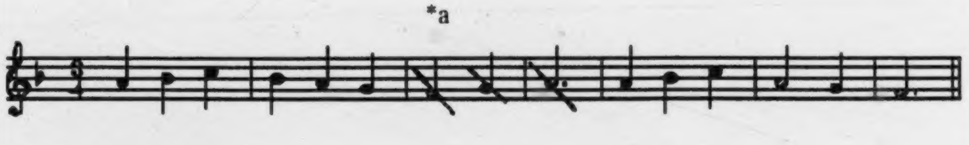
Indices of Difficulty and Discrimination. These items had difficulty indices of .53 and .51, respectively. Point-biserial indices of correlation for the correct response ($r_{pbis} = .44$ and $r_{pbis} = .34$) were within accepted guidelines.

Further analysis revealed that many students who provided incorrect responses had interpreted the items appropriately. Of those students who responded incorrectly to item 11, 5 students circled the note previous to the last note heard and 6 students circled the first note which followed the last note heard. Similarly, of those students who responded incorrectly to item 12, 3 students circled the note previous to the last note heard and 8 students circled the first note which followed the last note heard.

Item Revision. The number of responses which could not be categorized (6 and 4 for items 11 and 12, respectively) was problematic. However, considering that those students who answered the questions appeared to be following the music as requested, these items were not revised pending further information gathered from Field Study 2.

Items 13 and 14

Items 13 and 14 were designed to assess a student's ability to follow music notation. In these questions the aural stimulus was divided into three sections—the first and last were sung aloud and the second was silent. As illustrated in Figures 24 and 25, students were required to follow the notation and indicate which notes were silent (ie; the target notes).



[a]/ Indicates target notes.

Figure 24. Item 13: Target notes.

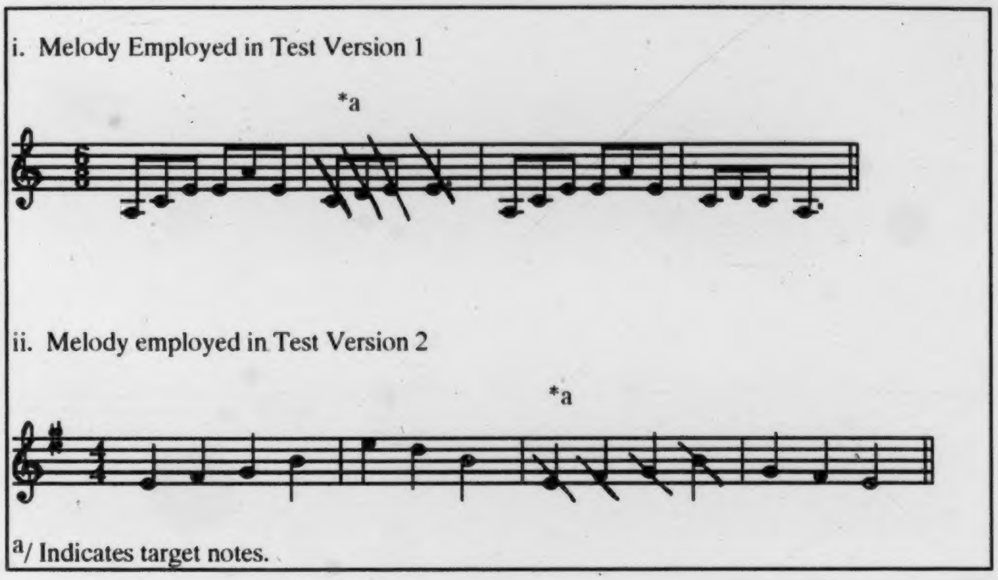

i. Melody Employed in Test Version 1
*a
ii. Melody employed in Test Version 2
*a
a/ Indicates target notes.

Figure 25. Item 14: Comparison of Test Versions 1 and 2.

Indices of Difficulty and Discrimination. Both items had a difficulty index of .06—well below the minimum of .20. Point-biserial correlations for the correct response, .33 and .50 respectively, were within accepted guidelines.

Of those who responded incorrectly, 13 and 20 students, respectively, identified the first note of the silent section. However, these students did not appear to know how to proceed from that point. Eight and 17 students combined the identification of the first silent note with the identification of up to four notes at the end of the melody. The remaining students who had identified the first silent note correctly proceeded to identify various other combinations of notes. Twenty-nine and 22 responses appeared random, indicating that students did not understand the questions. Four and 3 students omitted items 13 and 14, respectively.

Item Revision. A practise question was added in order to clarify these items. In addition, the melody in 6/8 meter employed in item 14 was replaced by a melody composed in 4/4 meter (see Figure 25). It was anticipated that use of a slower-paced melody (quarter note = 76 in 4/4 meter; dotted quarter note = 76 in 6/8 meter) with a smaller number of notes would facilitate accurate completion of the task.

Items 15 and 16

Items 15 and 16 were designed to assess a student's ability to identify the measure in which the aural presentation differed from the notational representation.

Indices of Difficulty and Discrimination. These items had difficulty levels of .55 and .63, respectively. Point-biserial correlations for the correct response, .24 and .29, respectively, were within accepted guidelines. However, the last measure of item 15 (r_{pbis} = -.07) and the first measure of item 16 (r_{pbis} = .00) were essentially nondiscriminating.

Item Revision. As presented in Test Version 1, the notation of both items 15 and 16 differed from the aural stimulus in the third measure. This may have provided a clue for the test-wise[2] students. In consideration of this, in Test Version 2, the notational representation of item 15 differed from the aural stimulus in the first measure (see Figure 26).

Further problems in item construction were revealed by the number of responses that could not be categorized—11 and 9 responses in items 15 and 16, respectively. Seven students indicated that, in both items, the aural presentation of the melody differed from the notation in more than one measure. Another 4 students made the same error in one of the two items. Consequently, the instructions were modified to inform students that the notation differed from the aural stimulus in one measure only.

Items 17 and 18

Items 17 and 18 were designed to assess a student's ability to identify the notational equivalent of an aural melody.

---

[2] Test-wiseness- "The ability to use assessment-taking strategies, clues from poorly written items, and experiences in taking assessments to improve one's score beyond what you would otherwise attain from mastery of the subject-matter itself" (Nitko, 1996, p. 124).

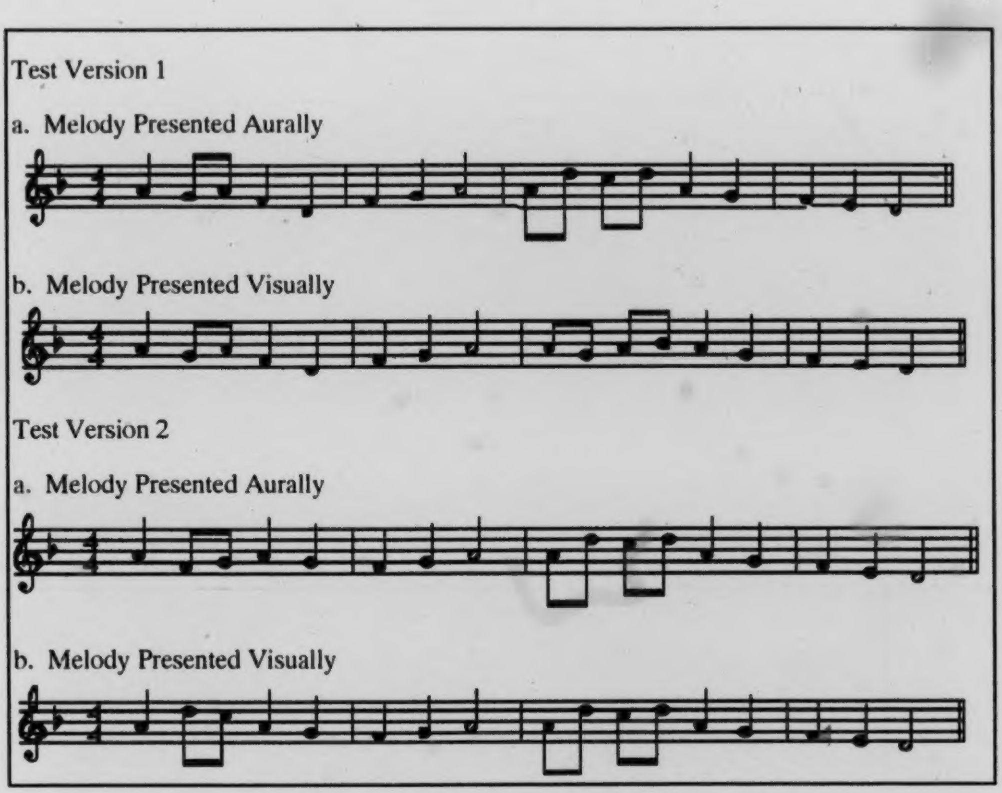

Test Version 1
a. Melody Presented Aurally
b. Melody Presented Visually
Test Version 2
a. Melody Presented Aurally
b. Melody Presented Visually

Figure 26. Item 15 : Comparison of Test Versions 1 and 2.

Indices of Difficulty and Discrimination. Difficulty indices were .47 and .27, respectively. Point-biserial correlations for the correct response were .47 and .36, respectively. In item 17, the point-biserial index of correlation for option B (-.14) was below the accepted guideline of .20. Option C discriminated acceptably ($r_{pbis} = -.35$). However, this option was chosen by only 14% of the students. In item 18, option B functioned properly ($r_{pbis} = -.31$); however, option A was essentially non-discriminating ($r_{pbis} = -.05$). As illustrated in Figure 27, both low-scoring and high-scoring students thought that the aural stimulus was represented by the descending step-wise pattern in the last two measures of distractor A rather than the disjunct motion notated in option C, the correct response.

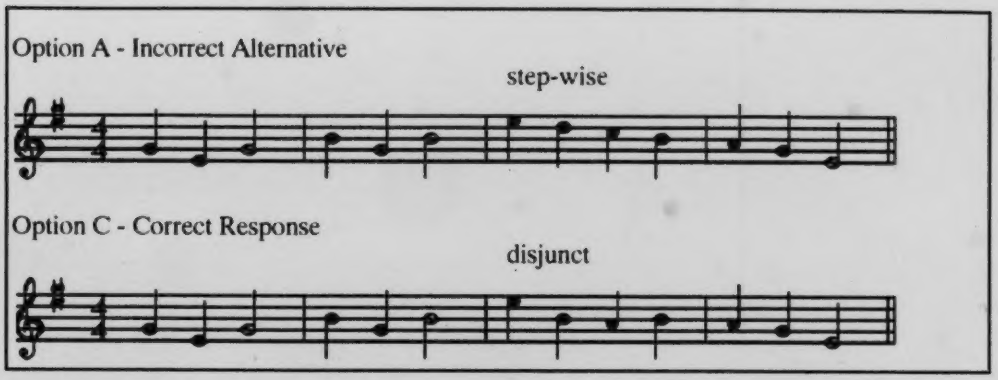

Option A - Incorrect Alternative
step-wise
Option C - Correct Response
disjunct

Figure 27. Item 18: Step-wise and disjunct motion.

Item Revision. In item 17, the second measure of option C was revised in order to make it more similar to the correct response (see Figure 28). Item 18 remained in its present form until further information was available to guide revision.

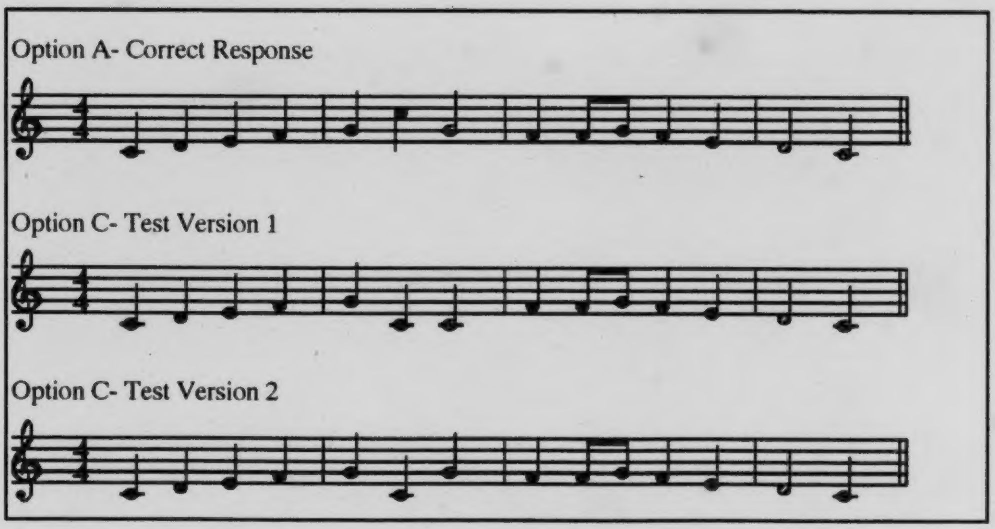

Option A- Correct Response
Option C- Test Version 1
Option C- Test Version 2

Figure 28. Item 17: Comparison of Test Versions 1 and 2.

Items 19 and 20

Items 19 and 20 were designed to assess a student's ability to determine the placement of the tonal center, commonly designated the home tone of a given melody.

Indices of Difficulty and Discrimination. These items had difficulty indices of .45 and .31, respectively. Point-biserial correlations for the correct response, .25 and .42 respectively, were within accepted guidelines.

Item 19, written in the key of 'a' minor, ended with the common A-G#-A configuration (see Figure 29). This was the most frequent response, being indicated by 22 of 49 students. The second most frequent response, indicated by 10 students, involved a step-wise descent to 'F'. The assumption that the melody descended at the end may be due, in part, to the familiarity of the descending mi-re-do motive in the major tonality. The note 'C', indicated by 5 students, functions as the home tone in the relative major tonality. There appeared to be no plausible explanations to clarify why the note 'G' was indicated by 1 student, the note 'B' was indicated by 2 students, or the note 'E' was indicated by 5 students. Responses from 3 students could not be categorized.

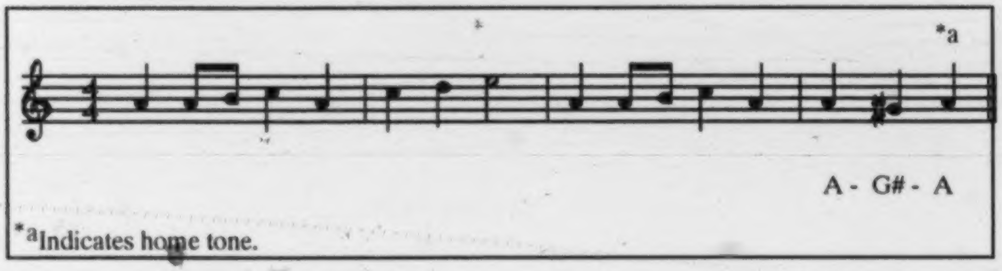


Figure 29. Item 19: Melodic material.

Although the melody composed for item 20 descended to the home tone (see Figure 30), either middle 'C' (indicated by 13 students) or high 'C' (indicated by 2 students) was accepted as the tonal center. While middle 'C' was the expected correct response, technically, the home tone is 'C'—the keynote of the scale. The 3 students who indicated 'G' may have assumed that the melody began and concluded on the same note. From a musical perspective, it was not expected that 13 students would identify the home-tone as 'A'. Perhaps this choice was influenced by the structure of the previous question which correctly ended on that note. No plausible explanations were uncovered to explain why the note 'B' was indicated by 2 students, the note 'D' was indicated by 2 students, the note 'E' was indicated by 4 students, or the note 'F' was indicated by 7 students. Responses from 3 students could not be categorized.

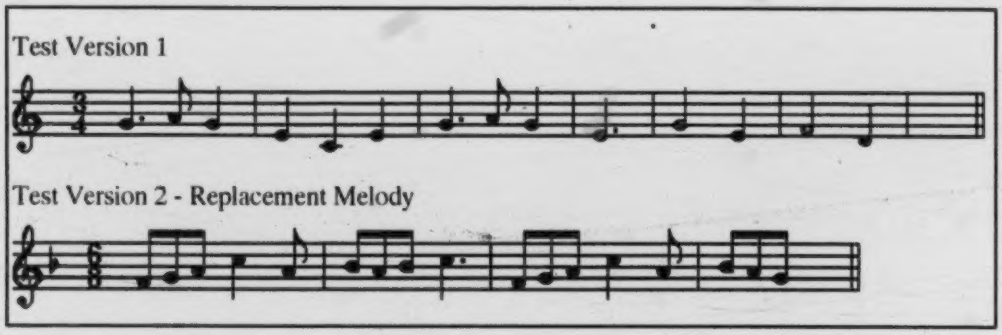


Figure 30. Item 20: Comparison of Test Versions 1 and 2.

Item Revision. For item 20, it was possible that students were uncomfortable concluding the melody on a ledger line note below the staff. Consequently, the melody was replaced by a melody in F-major which concluded with a step-wise descent to the tonic (see Figure 30). Item 19 was not revised.

Items 21 and 22

Items 21 and 22 were designed to assess a student's ability to identify a repeating melodic sequence. The melodies composed for each option began with the same melodic motive. Students were required to identify the melody which best continued the pattern in sequential fashion.

Indices of Difficulty and Discrimination. These items had difficulty indices of .37 and .47, respectively. A low point-biserial index of discrimination (.15) was reported for item 21. Similarly, indices of discrimination obtained for the incorrect options were below accepted guidelines (option B: $r_{pbis} = -.04$; option C: $r_{pbis} = -.04$). The internal structure of this item was examined to clarify why this item had not discriminated as expected.

Closer inspection revealed a structural flaw in the melody used in option C. The melodic sequence established in the first measure of each option was intended to be repeated in only option A, the correct response. The sequence established in the first measure of option C was repeated in the fourth measure of that option. Thus, as illustrated in Figure 31, both options A and C were correct choices. Although one could argue that option A best continued the sequence, the repetition of this pattern in the fourth measure of option C was an oversight in composition due, in part, to the difficulty of composing melodies which, within the framework of the question, sounded like real melodies rather than four-measure sequences developed for testing purposes.

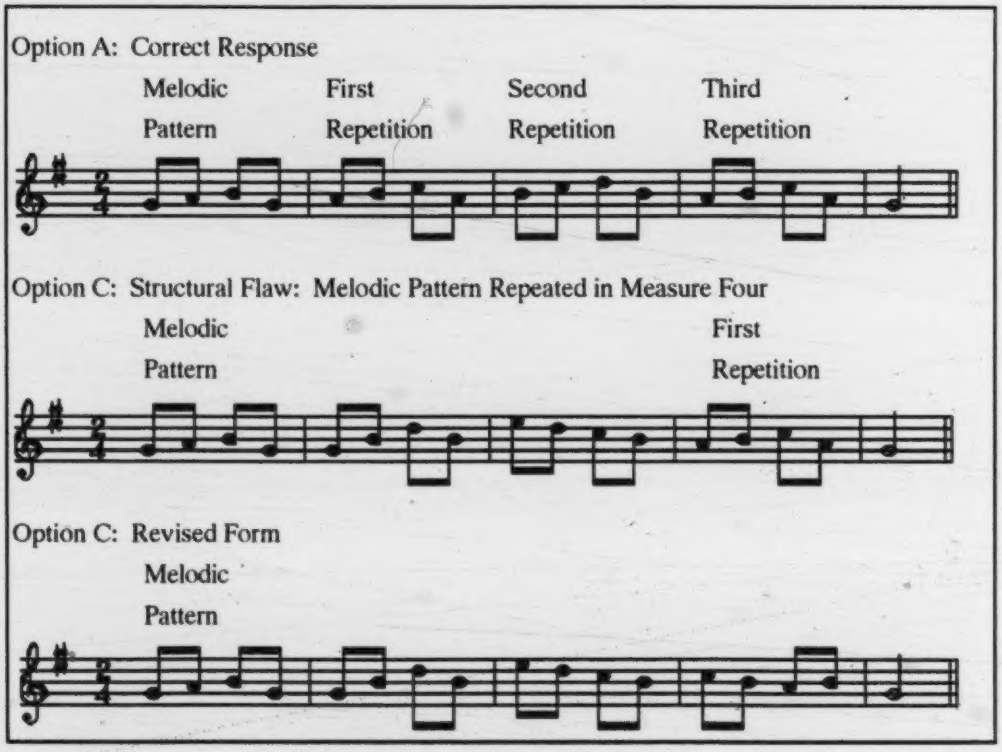

Option A: Correct Response
Melodic Pattern
First Repetition
Second Repetition
Third Repetition
Option C: Structural Flaw: Melodic Pattern Repeated in Measure Four
Melodic Pattern
First Repetition
Option C: Revised Form
Melodic Pattern

Figure 31. Item 21: Structural flaw revealed in option C.

For item 22, the correct response discriminated within the accepted guideline of .20 (r_{pbis} = .35). Incorrect options A and B were chosen by 18% and 33% of the students, respectively. Point-biserial indices of discrimination were below (option B: -.13) or approached (option A: -.22) the accepted guideline. Musically, no structural flaws were uncovered in this analysis.

Item Revision. In item 21, option C was revised so as to avoid the repetition noted above. No revisions were made to item 22 pending information obtained in Field Study 2.

Items 23 and 24

Items 23 and 24 were designed to assess a student's ability to identify the sharps or flats in an aural presentation of a melody.

Indices of Difficulty and Discrimination. Items 23 and 24 had difficulty indices of .18 and .41, respectively. Point-biserial correlations for the correct response, .27 and .42 respectively, were within accepted guidelines.

Point-biserial correlations were examined further to determine why item 23 was apparently more difficult than item 24. In item 23 the melody with no altered notes, chosen by 22 students, was a nondiscriminating option (option B, r_{pbis} = -.05). On the other hand, in item 24 the option with no altered notes (option C) was chosen by 21 students and had a point-biserial correlation of -.34. It is possible that, as a group, the items were susceptible to test-wiseness. High-scoring students were not sure how to respond to the first item in this set and thus chose the option with no altered tones. By the time they responded to the second item they realized that a note has been altered and therefore dismissed one of the options. This would account, in part, for the higher p-level in item 24.

In item 23, 16 students chose a flat when they heard a sharp. In item 24, 7 students chose a sharp when they heard a flat. In both cases, the options were essentially nondiscriminating (item 23, r_{pbis} = -.06; item 24, r_{pbis} = -.05).

Item Revision. Despite these apparent difficulties, these items remained in their present form until more data became available to guide revision.

Items 25 and 26

Items 25 and 26 were designed to assess a student's ability to notate an aurally-presented phrase.

Indices of Difficulty and Discrimination. Difficulty indices of .00 revealed that none of the students answered these items correctly. Twenty-nine and 34 students, respectively, provided written responses with the same number of notes as the aural presentation. Of these, 15 and 18 students submitted responses in which the contour of the written notes appeared to resemble the contour of the aural stimulus.

Item Revision. Item revisions, detailed in Figures 32 and 33, were intended to make the items easier. First, the melodies in 4/4 meter employed in Test Version 1 were replaced by motives written in 3/4 meter, thus reducing the number of unknown notes from 7 to 5. Second, the new melodies were based on familiar motives (a step-wise ascending pattern in a major tonality for item 25 and a minor tonic triad in an ascending and descending pattern for item 26).

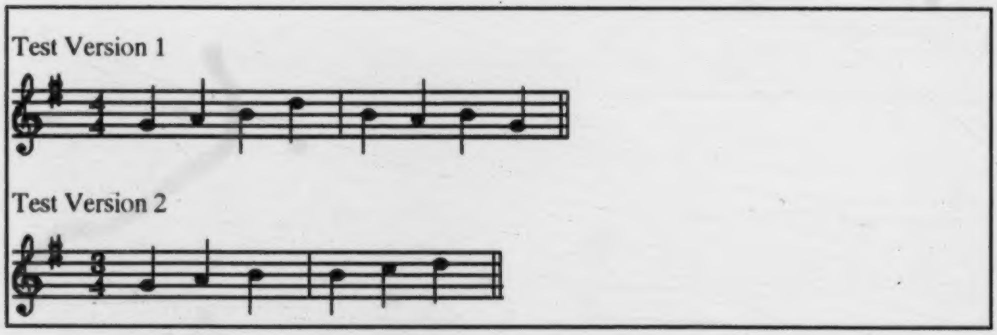


Figure 32. Item 25: Comparison of Test Versions 1 and 2.

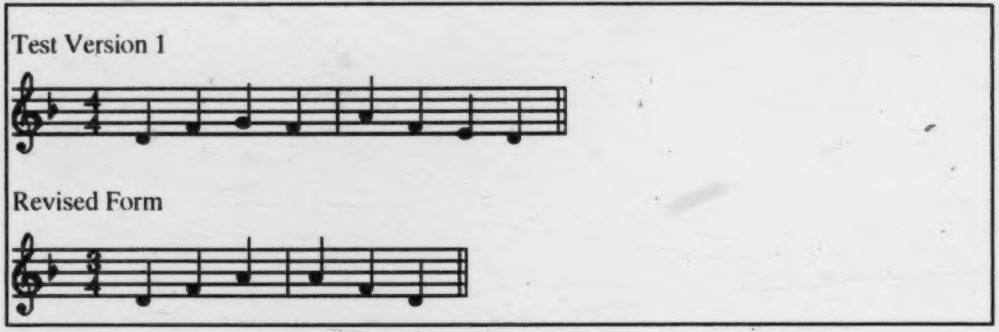


Figure 33. Item 26: Comparison of Test Versions 1 and 2.

Chapter Summary

Test Version 1, consisting of 26 items, was administered to students in two grade 5 classes in April, 1994. Forty-nine students completed the test. Item analysis was conducted in terms of classical test theory. The results of this analysis for each cluster or pair of items was described in Chapter 4. Based on this analysis, the diagrammatic representation of melody was changed to a representation in conventional music notation in items 1 and 2; melodies were modified in items 3, 5, 7, 14, 15, 17, 20, 21, 25, and 26; and directions to students were amended in items 5, 6, 15 and 16. Items 4, 8, 8, 9, 10, 11, 12, 13, 18, 19, 22, 23, and 24 were not revised. A practise question was added as a means by which to explain the requirement of items 13 and 14. A total of 26 items, which included the aforementioned revisions, formed Test Version 2 (see Appendix E), the administration and analysis of which is examined in Field Study 2 (see Chapter 5).

CHAPTER 5

FIELD STUDY 2: TEST VERSION 2 (ADMINISTERED INDIVIDUALLY)

Introduction

As described in Chapter 4, Test Version 1 was administered to 49 students in a group situation. Indices of difficulty and discrimination revealed in this analysis were used as a guide to item revision, the purpose of which, for the most part, was to make items easier. At this point in the study, it was evident that the process of item revision would be aided by information regarding the strategies employed by students while actually responding to the revised items. To obtain the necessary data, verbal reports were collected from students who were requested to think aloud while completing selected items from Test Version 2.

Method

Instrument

The length of time required for individual administration and the specific effects of a lengthy administration (see Millman & Greene, 1989) were unknown. It seemed clear, however, that not all 26 items could be administered. Therefore, for the purposes of Field Study 2, test items in each cluster were randomly assigned to two forms (Form A and Form B) so that each form contained 13 items. Form A was used in this administration. A copy of Form A is provided in Appendix E.

Student Sample

Participants were chosen in consultation with the music teachers in each of two schools located in School Division 1. The teachers of the grade 5 music classes in each school explained the study to their students and asked for volunteers. Eight students (four in each school) were selected at random from those who indicated they were interested in participating. Parental permission was obtained from these students before proceeding with the study (see Appendix C). Two of the 8 students were taking private piano lessons; the remaining 6 students received their formal music instruction solely within the school system.

Interviews

Each interview was completed in approximately 45 minutes. All verbal data, experimenter probes, and student protocols were recorded on a cassette tape. The transcription of this tape-recorded data yielded a "lean, spare description stripped of the contextual references in which the data originated" (Owens, 1982, p. 7). Thus, the credibility of the study was increased by allowing the use of verbatim accounts in the analysis, interpretation, and subsequent account of the student reports (LeCompte & Goetz, 1982).

The students were asked to think aloud while they responded to each of the musical questions in Form A. In preliminary introductions, students were told that their help was needed in gaining an understanding of how students in grade 5 responded to these test items (Scardamalia & Bereiter, 1983). This introduction included a brief summary of data-gathering procedures so that students understood why it was necessary to audio-tape their responses.

Assuming that students were not accustomed to describing their thoughts verbally while answering paper-and-pencil test items, the collection of data was preceded by two practise questions (Afflerbach & Johnson, 1984; Krutetskii, 1976), both of which were adapted from former studies (Ericsson & Simon, 1985; Krutetskii, 1976). The first, an addition question, was intended to obtain a concurrent report (Ericsson & Simon, 1980, 1985). The second, a description of their home-room classroom, was intended to evoke a retrospective report. These introductory questions were followed by a practise item chosen randomly from Form B. Structurally, this item was similar to item 7 in Form A.

Prior to the administration of Form A, students were instructed to:

> . . . say out loud everything that comes into your head that is related to the test. I am interested in what you are doing to figure out the answer, not just what the answer is. Do not try to explain anything to anyone else; pretend there is no one here but yourself.

One student became frustrated when he was unable to answer the questions immediately. Following attempts to answer the eighth question and given an opportunity to opt out of the study, this student chose not to continue. Thus, the protocol analysis is limited to the data obtained from the seven remaining students.

Analysis

The presentation and interpretation of these data are structured within Chapter 5 in two major sections. First, the data obtained in the study are examined through a detailed analysis of the experimenter's use of verbal prompts. Second, strategies employed by students, as evidenced in the student protocols, are reviewed on an item-by-item basis with specific reference to implications for item revision.

Analysis of Probes

Categorizations

Ericsson and Simon (1980, 1985) advocated the use of nondirected probes which encouraged students to keep talking. They maintained that probes in this form have a negligible effect on the student's internal processing of the information (Ericsson, 1987; Ericsson and Simon, 1985; Krutetskii, 1976). In order to comply with this contention, it was anticipated that probes used in the present study would be limited to general questions which asked the students what they were doing and general reminders to the students to keep talking. During the interview conducted with the first student, it became apparent that these probes did not sufficiently reveal the strategies employed by students. As a result, it was necessary to employ several types of probes in order to determine more completely the strategies employed by students.

Overall, probes were divided into five categories: open-ended reminders, requests for additional information, verbal cues for non-verbal behavior, attention to options, and implications for item revision. Each category is discussed in turn.

Open-ended Reminders

To illustrate the use of open-ended reminders, the use of open probes (p) with student 5 (s5) on item 10 occurred as follows:

Item 10

p- Can you tell me what you are thinking?

s5- Well it sounded like it [the note] was either 'F' or 'A' or a 'C' or a 'E'.

p- Okay. Anything else?

s5- No.

In addition, each task concluded with a dialogue similar to the following:

Item 2

p- Anything else you can tell me about that question?

s4- No.

p- Ready for the third one?

s4- Yes.

Non-directed comments related to the repetition of the aural stimulus were included in this category:

Item 1

p- Do you want to hear it one more time?

s3- Sure.

Requests for Additional Information

This category was divided into two subgroups of probes. The first subgroup of probes involved situations in which it appeared that students were using more knowledge to solve the tasks than what was revealed in the information stated overtly:

Item 8

s6- Well, I'm going back through the bars and the lines and I know that the two last bars with the lines over top were right and then the other one is right so it must be this one right here because it really didn't sound right to me.

p- Can you explain to me [about] the last two measures. You said that you knew those two are right. How do you know that they are right?

s6- Well, because her voice did the two titi's and went loo loo loo loo again and then it went down and then it went farther down.

The second sub-group involved probes in which students were requested to expand upon ideas which, from the researcher's perspective, appeared unclear (Norris, 1990; Patton, 1990; Scardamalia and Bereiter, 1983):

Item 2

p- What are you thinking as you do this?

s1- Well you go up one and up one.

p- What does up one mean?

s1- Well, if it's in a space you go to the next line.

Verbal Reminder

The third classification included probes used to ask students to explain aspects of the student's response which would not be recorded on the audio tape. For example, the use of this probe with student 6 on item 2 occurred as follows:

Item 2

p- *I see you have circled some notes.* Can you tell me what you were thinking?

s6- Well, when it said that I was supposed to circle the notes that went by step I looked at all of them and most of them had skips so I knew that it was just

the two because they were from 'A' to 'B' or from 'F' to 'G' and those were
the only notes.

In addition, students sometimes relied on visual information to supplement what they said. Within the total context the meaning was clear. However, to gain an aural record, a probe was used:

Practise Item

s6- Well, I was kind of confused because I don't think she said the first note in the bar.

p- Of option A?

s6- Of option A. And the same in the next bar in option A and basically everything else was the same.

In this case the student had pointed to option A.

Attention to Options

The probe classification, attention to options, referred to directed cues used in connection with the multiple-choice items. As illustrated below, after students had chosen an answer they were requested to explain why they had not selected the remaining options:

Item 9

s1- That ones pretty close to what she was singing.

p- Which one?

s1- [Option] A.

p- Anything else?

s1- No.

p- What about [options] B and C? Is there anything else you were thinking as you were listening?

s1- On [option] B she went up too high here and on [option] A she was right on—so it wasn't this one.

In defence of the use of these leading probes, Ericsson and Simon (1980) suggest that when researchers are interested in a certain aspect of the students' behavior the "verbal probe may be constructed to induce the students to general information specifically relevant to the hypothesis under consideration" (p. 222). In view of this, although the prompts may have directed students to attend to aspects of the question not formerly heeded, the information collected was essential in terms of item construction. Further, it appeared that attention to all options in an item was not beyond the capabilities of students at this grade level. This was exemplified by the behavior of student 2, who attended to all options without prompting:

Item 9

s2- Okay, I don't know if it's [option] A or [option] B, but I know it's not [option] C because she didn't go down to whatever that note is. When she got there she went up to [note] C.

Implications for Item Revision

The fifth classification, implications for item revision, related to statements supplied by the researcher in answer to questions of clarification from the students:

Item 3

s7- I have a question. I don't understand that.

p- Do, re, mi, fa, so, la, ti. Do you know the do, re, mi's?

s7- Now I understand. I can do that now.

In all cases these probes, preceded by requests for information on the part of the students, were intended to clarify directions or an item. The use of these probes suggested the need to revise the items.

Data Analysis

As described in Kail and Bisanz (1982), task analysis, as evidenced in verbal coding, depends upon the "theoretical framework of the investigator as well as the extent and sophistication of available research and theory" (p. 237). Although Ericsson and Simon (1980) advocated data analysis in terms of coding schemes defined a priori, they acknowledged that this may not be possible if cognitive tasks have not been investigated thoroughly in previous studies. Because the current study examined aspects of music literacy not assessed extensively in standardized music tests, coding guidelines were not established before the collection of data. Rather, they emerged during the coding process. This process was based, initially, on the professional judgement of the researcher.

Subsequently, the trustworthiness (Guba, 1981) of the data analysis was demonstrated by means of an external retrospective audit conducted in cooperation with a specialist music educator familiar with methods of naturalistic inquiry. Considering the number of protocols collected in this study (7 students x 13 test items), it was not possible to include all data collected in Field Study 2 in the external audit. Therefore, for practical purposes, items administered in Field Study 2 were divided into two groups: the first group was comprised of multiple-choice items (items 1, 15, 17, 21, and 24) and the second group was comprised of those items which required a written response (items 4, 5, 7, 9, 11, 13, 18, and 26). Two items were selected randomly from each group of items. Consequently, the external audit involved the examination of data obtained for items 1 and 15 (multiple-choice) and items 9 and 13 (written-response).

For purposes of the external audit, both the primary researcher and the external researcher reviewed the transcripts obtained for items 1, 9, 13, and 15. Following procedures described for data analysis, each researcher coded the data independently. Of a total of 127 coded sections, both researchers assigned 114 sections (90%) to the same category. This level of agreement supports the trustworthiness of the analysis conducted in this study.

In the present study, verbal reports were portioned into individual segments, each of which contained sufficient information for making an encoding decision (Ericsson & Simon, 1985). Coding categories were defined as clearly as possible, following Afflerbach and Johnson (1984) and Ericsson and Simon (1985). Due, in part, to the lack of previous research, categorizations were reviewed and revised throughout this process with special attention given to those segments which did not conform to the classifications (Afflerbach & Johnson, 1984).

The coding process treated each segment of the protocol separately, irrespective of the preceding or following segments (Ericsson & Simon, 1985). Coded segments were recorded on separate cards on which numbers written on the underside identified each student. Cards were shuffled prior to the categorization of each item. For the most part, categories were based on the musical properties of the responses.

Item-By-Item Analysis

The 13 test items administered in Field Study 2 were analyzed on an item-by-item basis. For each item, the analysis of data is presented in three sections: first, an analysis of responses from students who answered the item correctly; second, an analysis of responses from students who answered the item incorrectly; and third, where applicable, a description of item revisions. Item numbers refer to the order in which items were administered in Test Version 1.

Item 1

Item 1 was designed to assess a student's ability to identify the contour of a given melody.

Correct Response. Students 1, 3, 4, 5, 6, and 7 answered this item correctly. All students compared their description of the phrase they heard with the phrases presented in notational form. Students 1, 4, 5, and 7 described the contour of option B, the correct response. Student 1 stated, "[I chose option B because] she went up and up and then she went down and then she went down there." Students 6 and 7 compared correct and incorrect options. In the words of student 6, "I didn't think about number [option] A 'cause I know it wasn't that one . . . and I heard her voice going up but the choice C was going down so I knew that it would be the second question [option]." Similarly, student 7 eliminated option C, "[I didn't choose C because] you kept on going down, you didn't go up three notes and then go down one."

Student 3 did not describe the overall contour of the melody. Rather, he chose the correct option by comparing the position of the highest note in the aural phrase with the placement of that note in the notation. In this student's words, "There was also that high note. It was right here and that's not where you had it 'cause you had low down here, not high."

Incorrect Response. Student 2 based her response solely on the rhythmic structure of the melody. This was not a successful problem-solving strategy because all options were written in the same rhythm.

Item Revision. Students 1, 2, and 3 asked for clarification before completing this task. This indicated that directions to students were unclear. Directions to attend to contour were replaced by instructions to identify the option in which the written melody was the same as the melody they heard (see Figure 34).

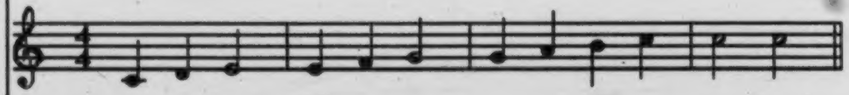
Test Version 2 - Directions to Students

Listen to the melody twice. Circle the letter above the melody in which the **overall shape** or **contour** is the same as the melody you hear.

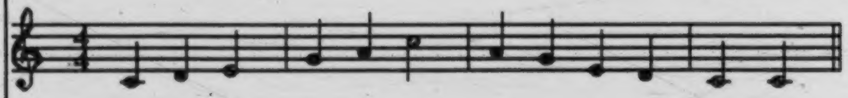
Test Version 3—Directions to Students

Listen to the melody twice. Circle the letter above the one melody which is the **same** as the melody you hear.

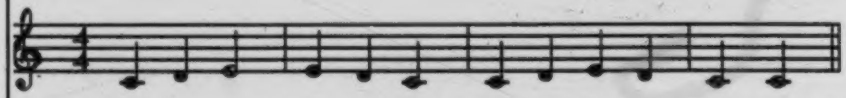
A



B



C



Note - Both Test Versions employ the same options.

Figure 34. Item 1: Comparison of Test Versions 2 and 3.

Item 4

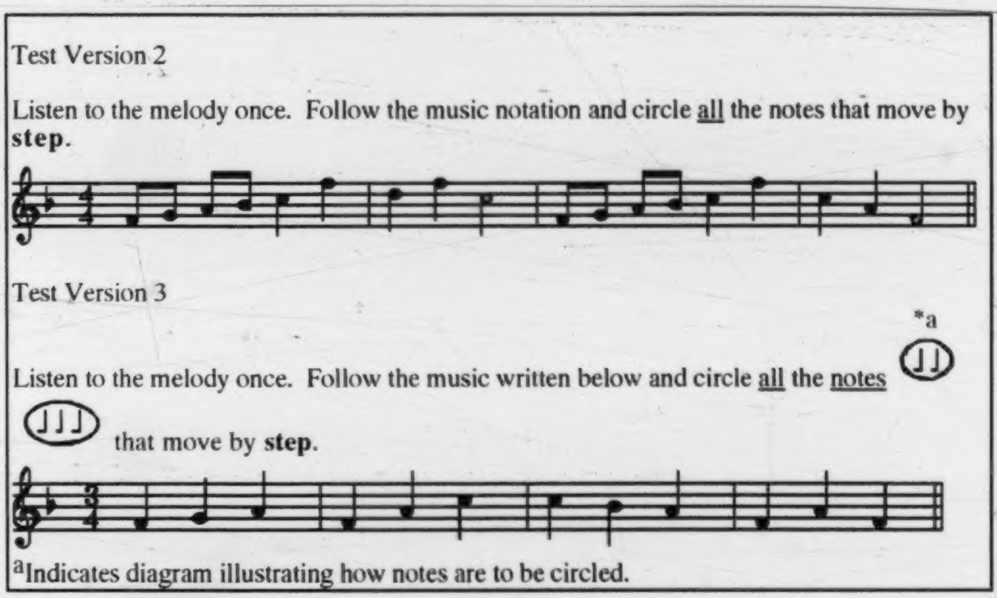
Item 4 was designed to assess a student's ability to identify all notes which moved by a step.

Correct Response. Students 1, 2, 3, 4, 6, and 7 answered item 4 correctly. Protocol analysis revealed that correct solutions were dependent, in part, upon defining and/or describing step-wise motion. Students 1, 2, 3, 4, and 7 identified the lines and spaces. Student 1 explained, "If it's in a space you go to the line." Students 3 and 6 referred to the letter names of the notes. In the words of student 3, "They were from 'A' to 'B' or from 'F' to 'G'." These strategies were accompanied by descriptions of leaps or skips. For example, in the words of student 7, "[There are] two spaces in between so that would be a skip." Variation in terminology did not influence the effectiveness of the strategy employed. Students 2, 3, and 4 extended their description of wider intervals to form direct comparisons. As student 4 explained, "If one's on a bar [referring to a line on the staff] and the other one isn't that means it's a step and it can't be a leap because a leap is more than two."

Incorrect Response. Student 5 answered this item incorrectly. This student combined a description of steps with a summary of the general contour, but did not refer specifically to leaps or skips.

Item Revision. Although the students did not refer overtly to rhythm, it appeared that the rhythmic structure of the melody influenced the responses provided by students 6 and 7. These students demonstrated their understanding of step-wise movement by circling the eighth notes that moved by step. However, in a response similar to that obtained in Test Version 1, these students did not circle the quarter note that completed this step-wise segment. Therefore, as illustrated in Figure 35, the melody employed in this item was replaced by a melody which contained only quarter notes. The melody employed for item 3 was revised similarly.

Student 3 did not know what was meant by 'notation'. This word was replaced by the phrase "music written below". Both students 4 and 5 asked how they were supposed to circle the notes. In consideration of this, the illustration of two or three quarter notes circled together, presented in the original test prototype but deleted at the request of one of the music teachers who reviewed the test, was reinstated (see Figure 35). For item 3, directions to students were changed similarly.



Test Version 2

Listen to the melody once. Follow the music notation and circle all the notes that move by **step**.

Test Version 3

*a

Listen to the melody once. Follow the music written below and circle all the notes that move by **step**.

[a]Indicates diagram illustrating how notes are to be circled.

Figure 35. Item 4: Comparison of Test Versions 2 and 3.

Item 5

Item 5 was designed to assess a student's ability to equate notes with their corresponding solfa syllables (do, re, mi, fa, so, la, ti).

Correct Response. Students 2, 6, and 7 answered this item correctly. These students related the order of the syllables to the order of the notes on the staff. Student 2 explained, "I take the order of do-re-mi-fa-so-la-ti-do. We start at 'do' and then the next note up would be 're'." All three students referred to previous notes as they proceeded with the question. In the words of student 7, "That's the same note so that would be the same." In addition, student 7 referred to the song Do Re Mi from the movie. The Sound of Music. Student 6 compared the solfa syllables to the absolute letter names, "If you start from the note 'C' it had a 'd' under it . . . then 'F' would be 'fa'."

Incorrect Response. Students 1, 3, 4, and 5 answered this item incorrectly. As revealed in the verbal protocols, students 1 and 4 were aware that the melody descended after 'mi' in the first measure. Nonetheless, this note was labelled 'fa' on the test papers. Student 4 explained, "Well, 'do', 're', 'mi' and then it started going down so I picked 'f' ['fa']." This error could be due, in part, to an inability to separate the order of the syllables when listed in succession from the order of the notes as they appeared in the given melody. Student 1 correctly identified 'so' and 'la' but continued an ascending line with the syllable 'ti' rather than falling back to 'so'. Student 3, who missed the skip from 'do' to 'so', named the notes in measure two fa-so-fa. His reference to the pentatonic scale, however, displayed a high level of analysis.

Student 5 based her decisions on her impression of the aural stimulus. This strategy was inadequate because her memory for the melody differed from the aural presentation and she was thus unable to describe the contour accurately.

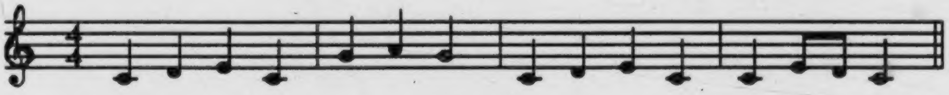
Item Revision. With the exception of student 3, all students requested clarification with regard to what was meant by d, r, m, f, s, l, t. During the interviews, when the examiner expanded these abbreviations to their full names, do-re-mi-fa-so-la-ti, students were able to proceed with the task. As illustrated in Figure 36, this revision was incorporated into Test Version 3.

Item 7

Item 7 was designed to assess a student's ability to identify notes by their absolute letter names (A, B, C, D, E, F, G).

Test Version 2 - Directions to Students
Listen to the melody once. Write the **solfa syllables** (d,r,m,f,s,l,t) on the lines under the notes.

Test Version 3 - Directions to Students
Listen to the melody once. Write the **solfa syllables** (do,re,mi,fa,so,la,ti) on the lines under the notes. The first one is given.



do _ _ _ _ _ _

Note - Both Test Versions employ the same melody.

Figure 36. Item 5: Comparison of Test Versions 2 and 3.

Correct Response. Students 1, 2, 3, 5, 6, and 7 answered this item correctly. Students 2, 3, and 7 used words or phrases to remember the names of the notes. The procedure was summarized by student 3, "The spaces are face F-A-C-E and the lines are Every Good Boy Deserves Fudge and I'm saying those in my head while I was going up." Students 1, 5, and 6 solved the task at what appeared to be an automatic level. Nonetheless, student 5 referred to memory devices when probed for more information.

Incorrect Response. Student 4 used memory devices similar to those employed by students 2, 3, 5 and 7, but was unable to apply them correctly.

Item Revision. Student 2 questioned the use of the phrase "absolute letter names". This expression was replaced by the phrase "names of the notes".

Item 9

Item 9 was designed to assess a student's ability to identify places in which notes in the melody were separated by a half step.

Correct Response. No correct responses were submitted.

Incorrect Response. All 7 students answered this item incorrectly. Comments from the students indicated that this aspect of melodic discrimination had not been formally introduced in their music classes. Students 1 and 4 outwardly stated that they did not know the meaning of the term half-step. In the first student's words, "Half step. What does that mean?"

Students 2, 3, and 7 related the half-step to the lines and spaces of the staff and correctly deduced that notes separated by a half step would be next to each other. Student 7 explained, "Those are circled because that goes from a line and to a space and back to a line." However, beyond this analysis, these students were unable to discriminate between half steps and whole steps. Students 1 and 6 associated half steps with melodic intervals in more general terms. In the first student's words, "These ones are too far away to be a half step." Only students 4 and 5 did not refer to half steps as the distance between the notes. The former defined half steps in terms of high notes and the latter in relation to rests. This revealed that, even though these students did not have the musical background necessary to complete the question correctly, they were not responding randomly.

Item Revision. It is possible that items dealing with the identification of semi-tones might belong more appropriately in a test of interval discrimination. Therefore, considering that this question posed considerable problems in both Field Studies 1 and 2, both items which assessed this aspect of melodic perception were deleted from the test.

Item 11

Item 11 was designed to assess a student's ability to follow music notation. In this question only part of a given melody was sung aloud. Students were required to circle the final note they heard.

Correct Response. This item was answered correctly by students 1, 2, 3, 4, 5, and 6. Five students prefaced their response with what seemed to be an instant identification of the correct note. Student 5 explained, "She stops singing at that place." Further explanation related to melodic, metric, and rhythmic aspects of the melody.

Students 1, 4, and 6 used the contour as a guide to determine the note on which the aural presentation ended. In the first student's words, "She sang the first two bars which went up and it stayed the same there and then went down and then it stayed on that note." Student 1 depended solely on this strategy. Student 4, who depended mostly upon contour in response to this item, alluded to meter or rhythm by stating, "She only used that sort of amount of notes." In this case, it is possible that this student counted the number of beats she heard. In contrast, student 6 began her analysis by integrating contour and rhythm, "[She sang] the one note and then she went higher with the 'ti-ti'." After analyzing the first three notes in this manner, she relied solely on the contour.

Students 4 and 6 demonstrated the use of contour as the primary strategy and rhythm as a secondary strategy with the response guided mainly by the former. On the other hand, students 2 and 3 attended solely to metric or rhythmic aspects of the melody. Student 2 referred to the duration of half and quarter notes but did not explain how this information was used beyond stating that, "[It] helps to follow [the notes]."

Incorrect Response. Student 7 answered item 11 incorrectly. This student aurally divided the melody into two equal phrases and described how the singer paused for a breath after the second measure before singing the final note. Interestingly, after describing the aural presentation correctly, she circled the wrong note. The reason for this error cannot be inferred from the data.

Item Revision. No students asked for clarification before completing this task. Therefore, the item was not revised.

Item 13

Item 13 was designed to assess a student's ability to follow music notation. In this question the aural stimulus was divided into three sections—the first and last were sung aloud and the second was silent. Students were required to follow the notation and indicate which notes were not sung aloud.

Correct Response. Students 1, 2, 3, 6, and 7 answered this item correctly. Students 6 and 7 referred specifically to the contour of the melody. In the words of student 6, "She went low and then stopped and then she went to [a] high [note] and then down to [a] low [note]." Students 1, 2, 3, and 7 examined the rhythmic structure. Student 2 referred to the "timing of each note." Student 3 identified the second section by counting the number of silent beats, "You stopped here for six beats and then you just kept going." Students 1 and 7 extended this strategy with reference to the duration of the dotted half note which denoted the end of the second section. In the words of student 1, "she started up in quarter notes after the [dotted] half note."

Incorrect Response. This item was answered incorrectly by students 4 and 5. Student 4 was unable to identify the notes which comprised the second section because she concentrated exclusively on the final section of the melody. Student 5, who referred to the duration of the dotted half note which denoted the end of the second section, was unable to correctly identify the silent notes because her memory of the aural stimulus differed from the actual presentation. This illustrated that the accuracy of the response did not depend entirely upon the detail with which the rhythm was analyzed.

Item Revision. Five students answered this question correctly. In comparison, this question had a difficulty index of .06 in the pilot study which was well below the generally accepted guideline of .20. The apparent ease with which students in the protocol sample completed this task may be due, in part, to the inclusion of a practise question. The data did not reveal the need for further revision.

Item 15

Item 15 was designed to assess a student's ability to identify the measure of a given melody in which the aural presentation differed from the notational representation.

Correct Response. Students 2, 3, 6, and 7 answered this item correctly. Students 3 and 7 compared the contour of the aural and notational presentations with specific reference to the measure in which they differed. In the third student's words, "They [the notes] went high instead of low." Student 7 illustrated this difference by singing both versions. Student 6 examined the melody beginning with the last measure and moving forward to the first. By employing this order of analysis, the correct response was chosen through the process of elimination. Student 2 based her response on the rhythmic structure of the melody, even though the rhythm of the aural and notational versions did not differ. It was not known how this student was able to discern the correct response by employing this strategy. It is possible, however, that she employed additional strategies which were not verbalized in the protocol and which were not exposed by the use of probes.

Incorrect Response. Incorrect responses were recorded by students 1, 4, and 5. Student 1 compared the notation to the aural version as she remembered it and decided that the first three of four measures were different. It appears that this student did not consider the directions in which she was instructed to choose one measure. Students 4 and 5 attended to the contour in general terms. "It looked like the right shape," said student 4. For the most part, these students concentrated on the rhythmic structure. As explained by student 5, "The first one sounded a little different too. It sounded like there was a 'ti-ti' in there." Student 4 did not accurately remember the melody as presented on the tape, "[She sang] only two notes on the second measure."

Item Revision. The one student who requested clarification asked if there was only one correct answer even though this was indicated in the directions with bold type. In the revised version emphasis on the word one was accentuated further by being underlined. In addition, the size of the font for this word was increased from 12-point to 14-point.

Item 17

Item 17 was designed to assess a student's ability to identify the notational equivalent of an aural melody.

Correct Response. Option A, the correct response, was chosen by students 1, 2, 3, 4, and 7. Student 4 described the overall contour of the correct option, "The shape was sort of like what she sang it." This response was extended with specific reference to the measure in which option B differed from the aural stimulus, "[It's not B because] right here it's really high." Student 1, on the other hand, described option A in terms of the third option, "[Option] C was low when it was supposed to be high. [Option] A was high at that point." Similarly, student 2 considered option A in terms of option B, "You didn't go up in the third measure." Moreover, both students 1 and 2 described the differences between the remaining option and the aural presentation of the melody. Student 7 chose the correct option by eliminating incorrect alternatives, "I know it didn't go down so it can't be [option] C." Student 3 used a similar strategy.

Incorrect Response. Two incorrect responses were recorded for this item. Student 5 chose option C because, in that option, the notation represented the melody as she remembered it. Options A and B were eliminated because they differed from this impression. Student 6 compared options B and C and correctly identified the note on which the latter option differed from the aural presentation. However, she neglected to analyze option A. This oversight most likely contributed to the error in response.

Item Revision. No students asked for clarification. Therefore, this item was not revised.

Item 19

Item 19 was designed to assess a student's ability to determine the placement of the tonal center, commonly designated as the home tone of a given melody.

Correct Response. Students 3, 5, 6, and 7 answered this item correctly. In school music classes, students are often taught that melodies begin and end on the home tone. Although, technically, this is not the way in which all music is constructed, test item 19 satisfied this definition. Using this rule, students 3 and 6 placed the final note on the same space as the note on which the melody began. Student 3 explained, "[You take] the first note and put it in [the] last because that's the home tone." This student checked his answer aurally by humming the last note when the melody was presented a second time.

Student 5 combined attention to the sound of the melody with reference to the letter names of the notes, "It sounded like it went just up a bit so I thought it would be 'A'." Similarly, student 7 sang the final note and described where it was placed in terms of note names, "I just think it [would] sound right going from an 'A' down to a 'G#' and then back to 'A'."

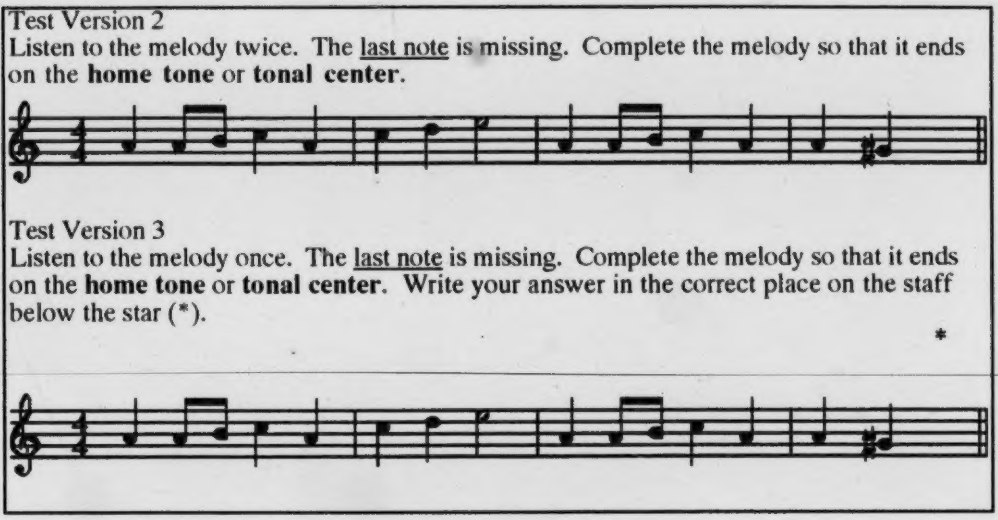
Incorrect Response. Students 1, 2, and 4 answered this item incorrectly. Student 1 employed an incorrect version of the home tone rule, "It will always go down low for a home tone at the end of the song." Student 2 compared the rhythmic structure of the aural stimulus to that of the notation. Student 4, who concentrated on the aural stimulus, justified his response by describing the contour, "On the last bar she started going down and I think it would sound good if it went down."

Item Revision. Students 4 and 7 did not understand the instructions. Upon further review it became evident that, although students were informed that the last note was missing, instructions were not prescriptive. As illustrated in Figure 37, in the revised version students were requested to write their answer on the staff in the area below the printed star.

Item 21

Item 21 was designed to assess a student's ability to identify a repeating melodic sequence.

Correct Response. The item was answered correctly by students 1, 5, and 6—all of whom observed the contour of the sequential motive in relation to the contour of the motives as presented in each measure. These students focussed specifically on notes in which the contour was the same or different. Student 6 explained, "In [option] A and



Test Version 2
Listen to the melody twice. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**.

Test Version 3
Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).

Figure 37. Item 19: Comparison of Test Versions 2 and 3.

[option] B they didn't repeat because . . . the first line was the same as [option] C and then she went from high to low and then low to high." Student 6 extended this response through reference to the rhythmic structure, "I'm gonna choose letter [option] C because . . . she went from low to high in her two 'ti's' and her second 'ti-ti' was high to low." Student 5, who correctly identified the sequence repeated in option C, further justified this response, "I think that one sounded the most like the one at the beginning."

Incorrect Response. Students 2, 3, 4, and 7 answered this item incorrectly. These students observed the overall contour, range, or rhythm of the melodies presented in each option rather than those aspects as they related to the one-measure sequence. Students 2 and 3 eliminated option C because the range of the melody differed from that of the other two options. In the words of student 3, "I don't think that sharp and that ledger line really fits here." Students 2 and 7 chose option B because it had the proper rhythmic structure. Student 4 based her preference for option B on the overall contour. Student 3, on the other hand, rejected this option because the melodic line descended consistently in the last two measures. He chose option A because the range of the notes was similar to that of the motive presented in the first measure.

Item Revision. Students 2 and 7 asked the examiner to clarify directions before completing the task. In consideration of this, modified directions explained more specifically how the options differed (see Figure 38). Furthermore, as illustrated in Figure 39, the melody composed for option C was revised in order to avoid the sharp and ledger line. The range of the new melody resembled more closely that of the incorrect options.

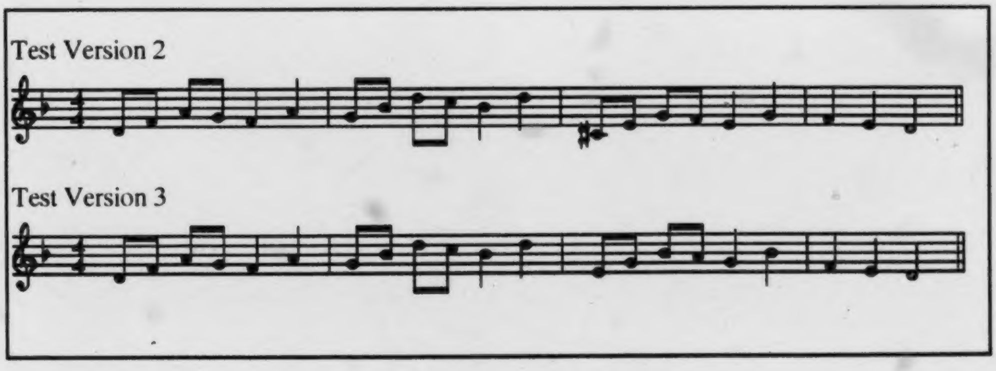
Test Version 2

Listen to each melody once. The melodies all begin with the same melodic **pattern** or **sequence**. Circle the letter above the music notation which **best continues** the repeating melody.

Test Version 3

Listen to each melody once. The first measure of each melody below begins with the same melodic **pattern** or **sequence**. Measures 2 and 3 are different. Circle the letter above the **one** melody which **best continues** the repeating melody.

Figure 38. Item 21: Directions to students.


Test Version 2

Test Version 3

Figure 39. Item 21: Comparison of Test Versions 2 and 3 for option C.

Item 24

Item 24 was designed to assess a student's ability to identify the sharps or flats in an aural presentation of a melody.

Correct Response. Students 4 and 7 answered this item correctly. Student 4, who concentrated on the overall contour of the melody as presented in option A, did not refer overtly to the sharps and/or flats. Student 7 related the aural stimulus to her perception of how sharps and flats influence tones. In this student's words, "I didn't think it sounded like an E sharp . . . I think it sounded mostly like an E flat."

Incorrect Response. Incorrect responses were submitted by students 1, 2, 3, 5, and 6. Student 2 correctly noted that the tonal patterns in each option were identical, "The notes are the same here except there's no sharps and flats." Student 1 provided a similar response. Students 3, 4, and 6 did not overtly describe the tonal patterns of each option. It is possible, however, that their observance of the sharps and/or flats was accompanied by non-verbalized attention to similarities among the tonal patterns presented in each option.

Students 2, 3, 5, and 6 related the aural stimulus to their perception of how sharps and flats influence tones. Student 5 aptly described this function, "This one [E flat in option A] seemed a little bit deeper and this one [E sharp in option B] seems higher." Students 2, 3 and 5 chose option C because, in their view, the melody was unaltered by either flats or sharps. In the words of student 3, "It sounded about the same as a regular 'E'." In contrast, students 1 and 6 realized that the tones had been altered. However, these two students thought that the aural representation of the melody contained sharps.

Item Revision. The apparent confusion with this task could be due, in part, to the difficulty of the question. Therefore, in Test Version 3, students were directed to listen for the sharps or flats (see Figure 40). In addition, the word "one" was accentuated with bold print and an underline.

Test Version 2- Directions to Students

Listen to the melody twice. Circle the letter above the one melody which is the **same** as the melody you hear.

Test Version 3- Directions to Students

Listen to the melody twice. Listen for the sharps or flats and circle the letter above the **one** melody which is the **same** as the melody you hear.

Figure 40. Item 24: Comparison of Test Versions 2 and 3.

Item 26

Item 26 was designed to assess a student's ability to write the notational equivalent of an aurally dictated phrase.

Correct Response. Students 3 and 6 answered this item correctly. Student 3 described the overall contour, "it sort of went about two notes high and this went up even higher about two notes higher and then it went low and then it went even lower." This student decided upon the final note by applying the home tone rule, "I think it would end on the home tone . . . and that one's a home tone." The final three notes were determined from this point of reference. Student 6 described the intervals between the notes, "I think that she skipped some notes so I would skip one note, draw one note and then skip another note and draw another note." The mirror-image of the melody (see Figure 41) was described by student 6, "I put down the same notes except going downward because it sounds the same as the ones that were going up."

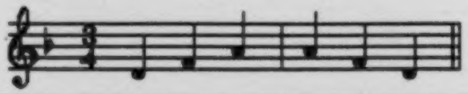


Figure 41. Item 26: Melody employed in dictation task.

Incorrect Response. Students 1, 2, 4, 5, and 7 answered this item incorrectly. Students 1, 4, and 7 realized that the second measure was the mirror-image of the first.

Errors were therefore attributed to the failure to distinguish individual intervals in the ascending portion of the melody. Student 7 perceived the given triad as a step-wise pattern. Student 1 perceived the first two intervals as a perfect fifth followed by a major second. The size of this interval may explain why she did not end the melody on the tonic. Even though student 4 did not refer specifically to the size of the intervals, she correctly identified the first two notes, but proceeded from there up a major second rather than a major third. The response supplied by student 5 differed from the aural stimulus to the greatest extent. Her impression of the minor triad began with the interval of a major sixth followed by a minor seventh. In addition, this student did not observe the relationship between the first and second measures. More specifically, from the end of the first measure her melody descended a minor second to 'A', ascended a minor third to 'C', and then descended a perfect fourth to 'G'. Although the student based the response on the contour of the given melody, this version did not reflect the overall shape of the original. Student 2 correctly identified the first four notes but did not end the melody on the home tone.

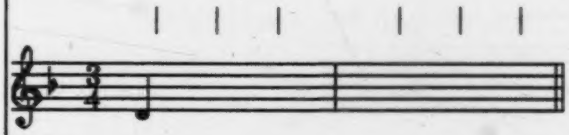
Item Revision. Students 2 and 6 indicated that they did not know how to write the rhythm of the melody. As illustrated in Figure 42, this aspect of the question was clarified.

Chapter Summary

Verbal protocols were obtained from 7 students who were requested to think aloud while completing 13 test items, 1 of which was referenced to each of the 13 instructional objectives identified in Chapter 3. For each item, the analysis of data was presented in three sections; first, an analysis of responses from students who answered the item correctly, second, an analysis of responses from students who answered the item incorrectly, and third, where applicable, a description of item revisions. Based on this data, directions to students were amended in items 1, 4, 5, 7, 15, 19, 21, 24, and 26; melodies were modified in items 4 and 21; and item 9 was deleted from the study. Items

Test Version 2

Listen to the melody three times. Write the melody on the staff below. The first note and the rhythm are given.



Test Versions 3

Listen to the melody three times. Write the notes of the melody on the staff below. The rhythm is shown above the melody in quarter notes. The first note of the melody is given.

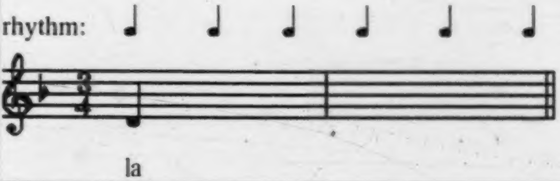


Figure 42. Item 26: Comparison of Test Versions 2 and 3.

11, 13, and 17 were not revised. Trustworthiness of the data analysis was demonstrated by means of an external retrospective audit. The revised items formed Test Version 3 (see Appendix F), the administration of which is examined in Field Study 3 (see Chapter 6).

CHAPTER 6

FIELD STUDY 3: TEST VERSION 3 (GROUP ADMINISTRATION)

Overview

The purpose of Field Study 3 was to examine the functioning of test items assigned to Test Version 3 when they were administered to groups of students. To this end, Test Version 3 was administered to 77 students in four grade 6 classes in October, 1994. All students attended one of four schools within School District 1 or School District 2 and each class was taught by one of four music specialists. One teacher had participated in Field Study 2. Within this class, the students who had completed the protocol analysis did not participate in the group administration. Each student who participated in Field Study 3 returned a form signed by a parent or guardian indicating that the student was allowed to take part in this study.

Test Versions 1 and 2 were administered to grade 5 students towards the end of the school year (April, 1994). Test Version 3 was administered to grade 6 students during the beginning of the school year (October, 1994). It was anticipated that this difference would not influence performance.[1]

As compared to Test Versions 1 and 2, the order in which the items were presented in Test Version 3 was revised considerably. Previously, items had been arranged according to the overall expected level of difficulty regardless of the item format (multiple-choice, written response) of the individual items. In contrast, in Test Version 3 items were ordered according to the format in which they were written. Test Version 3 began with those items whose structure was demonstrated by means of a practise question. This included tasks in which students drew a line across the final note sung aloud (items 1 and 2) and tasks in which students drew lines across the notes that were not sung aloud (items 3 and 4). These tasks were followed: first, by the remaining items in which students were required to supply a written response (items 5 through 14) and second, by the items in which students were required to choose their response from either three options (items 17 through 24) or four options (items 15 and 16). As discussed in Chapter 5 (see page 121) one pair of items referenced to an instructional objective identified in the analysis of selected educational sources was deleted from the study. Thus, in Test Version 3, a total of 24 items were referenced to 12 instructional objectives. A copy of Test Version 3 is provided in Appendix F.

---

[1] Altogether, six field studies were conducted during the development of the test items. It was desired that the data in each study be completed at the same time during the school year, namely toward the end (April) of the school year. This was not feasible given the time constraints. Consequently, the students included in Field Study 3 were assessed at the beginning (October) of the school year. This procedure is modelled after norming procedures formed with standardized testing. The assumption is that performance of the students at the beginning of a grade is not unlike the performance of students at the end of the previous grade on a similar test. Comparison of corresponding difficulty indices and point-biserials on the items that were essentially not changed during the duration of the study among Field Studies 1, 3, 5, and 6 reveals differences attributable to sampling error. Consequently, the assumption is made tenable.

The presentation and interpretation of data obtained in Field Study 3 is structured within this chapter in two main sections: first, indices of difficulty and discrimination are discussed with particular emphasis on implications for item revision; and, second, summary test statistics are used to clarify the structure of the items when viewed as an aggregate.

Item Revisions

The difficulty indices and point-biserial correlations for the correct response and incorrect options are reported in Table 18. These indices were analyzed with particular emphasis on implications for item revision. As shown in Table 18, the difficulty indices and point-biserial correlations for both the correct and incorrect options meet the criteria set for these students for 9 items: 3, 7, 8, 9, 10, 11, 12, 21, and 22. Consequently, no changes were made for these items. The corresponding indices for the remaining 13 items were not totally satisfactory. Consequently, changes were made to items 1, 2, 4, 5, 6, 15,

Table 18

Field Study 3: Indices of Difficulty and Discrimination

Item	Task	Format	Index	Options A	B	C	D	Revise
1 (11)	follow notes	WR	p	.84*	.16			x
			r_{pbis}	.39	-.39			
2 (12)	follow notes	WR	p	.83*	.17			x
			r_{pbis}	.40	-.40			
3 (13)	follow notes	WR	p	.47*	.53			
			r_{pbis}	.41	-.41			
4 (14)	follow notes	WR	p	.73*	.27			x
			r_{pbis}	.41	-.41			
5 (3)	leap (repeat)	WR	p	.18*	.82			x
			r_{pbis}	.25	-.25			
6 (4)	step	WR	p	.13*	.87			x
			r_{pbis}	.29	-.29			
7 (7)	names	WR	p	.40*	.60			
			r_{pbis}	.55	-.55			
8 (8)	names	WR	p	.65*	.35			
			r_{pbis}	.63	-.63			
9 (5)	solfa	WR	p	.18*	.82			
			r_{pbis}	.47	-.47			
10 (6)	solfa	WR	p	.45*	.55			
			r_{pbis}	.59	-.59			
11 (19)	home tone	WR	p	.45*	.55			
			r_{pbis}	.40	-.40			
12 (20)	home tone	WR	p	.40*	.60			
			r_{pbis}	.57	-.57			
13 (25)	dictation	WR	p	.12*	.88			
			r_{pbis}	.29	-.29			

Note. * denotes correct response. MC = multiple-choice; WR = written response.

Numbers in parentheses refer to item order in Field Study 1.

Table 18 (continued)

Field Study 3: Indices of Difficulty and Discrimination

				Options				
Item	Task	Format	Index	A	B	C	D	Revise
14 (26)	dictation	WR	p	.22*	.78			
			r_{pbis}	.31	-.31			
15 (15)	read notes	MC	p	.60*	.14	.17	.09	x
			r_{pbis}	.46	-.33	-.19	-.15	
16 (16)	read notes	MC	p	.04	.12	.78*	.06	x
			r_{pbis}	-.09	-.36	.41	-.14	
17 (1)	contour	MC	p	.05	.94*	.01		x
			r_{pbis}	-.14	.18	-.12		
18 (2)	contour	MC	p	.08	.61*	.31		x
			r_{pbis}	-.27	.41	-.28		
19 (17)	read notes	MC	p	.83*	.13	.04		x
			r_{pbis}	.37	-.40	-.03		
20 (18)	read notes	MC	p	.43	.16	.42*		x
			r_{pbis}	-.19	-.10	.26		
21 (23)	sharps	MC	p	.25	.27	.48*		
			r_{pbis}	-.25	-.07	.28		
22 (24)	flats	MC	p	.49*	.32	.18		
			r_{pbis}	.18	-.08	-.14		
23 (21)	sequence	MC	p	.18	.35	.45*		x
			r_{pbis}	.00	-.09	.11		
24 (22)	sequence	MC	p	.55*	.13	.30		x
			r_{pbis}	.39	-.15	-.26		

Note. * denotes correct response. MC = multiple-choice; WR = written response.

Numbers in parentheses refer to item order in Field Study 1.

16, 17, 18, 19, 20, 23, and 24. A practise question was added in an attempt to ensure that students know how to proceed with items 13 and 14.

The item-by-item analysis which follows is limited to a description of those items which were revised in consideration of the data obtained in Field Study 3. In this respect, results are discussed below for each cluster or pair of items. The exception is test item 4, which is discussed in the absence of item 3 (an item which was not revised). Unless specified otherwise, item numbers refer to the order in which items were administered in Test Version 3. Numbers in parentheses refer to the order in which the items were administered in Field Study 1. The analysis of the items not revised in Field Study 3 is provided in Appendix G.

Items 1 and 2 (Items 11 and 12 in Field Study 1)

Items 1 and 2 were designed to assess a student's ability to follow music notation. In these questions, only part of the melody was sung aloud. Students were required to draw a line through the final note they heard.

Indices of Difficulty and Discrimination. These items had difficulty indices of .84 and .83, respectively; both items discriminated positively (r_{pbis} = .39 and .40).

Item Revision. The difficulty indices obtained for items 1 and 2 were above the guideline of .80 generally accepted for this type of measure. In order to make item 1 more difficult, the simple rhythmic structure of the melody was revised to include eighth notes (see Figure 43). Item 2 was made more difficult by placing the target note on a musically less important beat or portion thereof. Therefore, in item 2, the target note on beat 4 in 6/8 meter was replaced by a target note on beat 6 in 6/8 meter (see Figure 44).

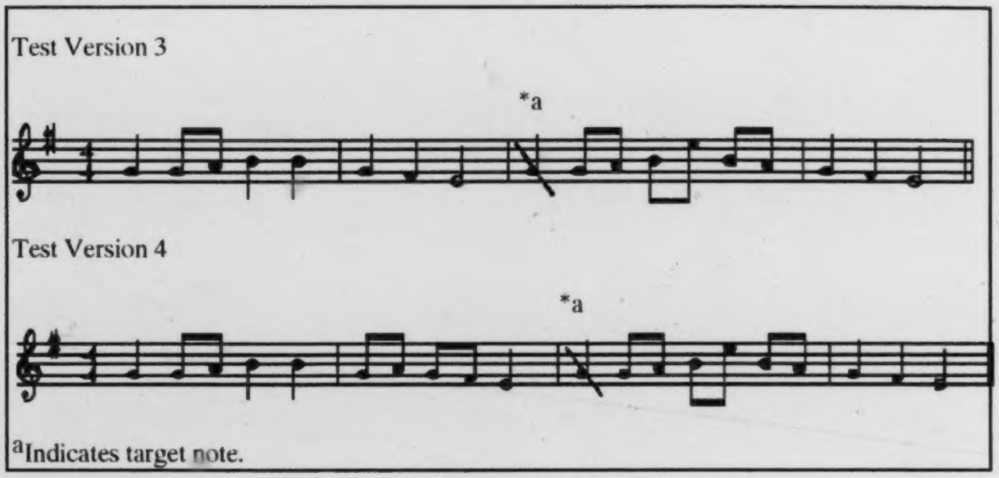

Test Version 3
*a
Test Version 4
*a
aIndicates target note.

Figure 43. Item 1: Comparison of Test Versions 3 and 4.

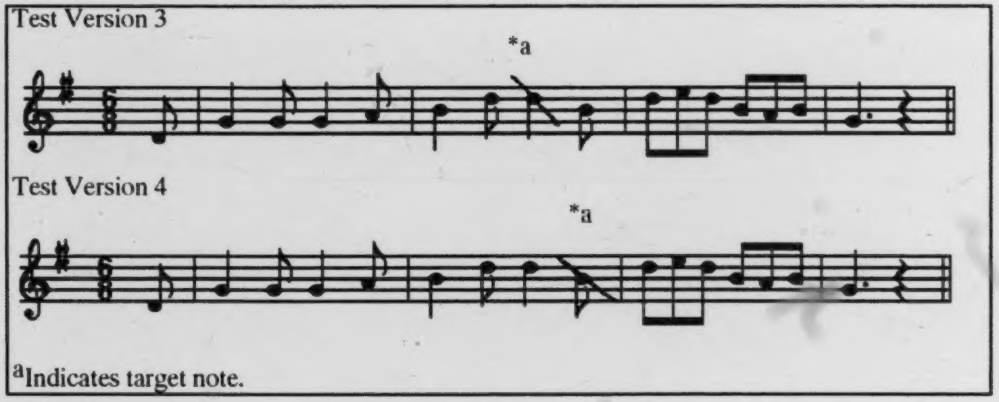
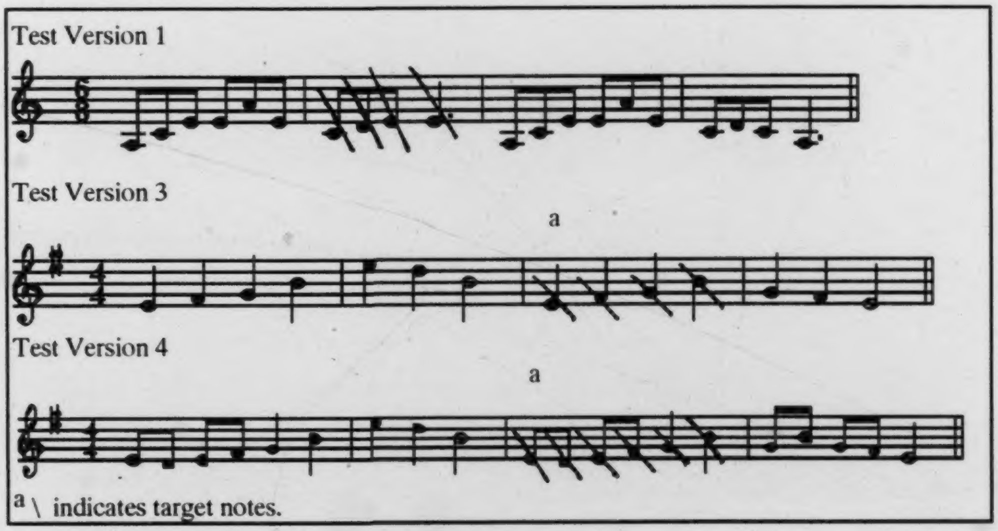

Test Version 3
*a
Test Version 4
*a
aIndicates target note.

Figure 44. Item 2: Comparison of Test Versions 3 and 4.

Item 4 (Item 14 in Field Study 1)

Item 4 was designed to assess a student's ability to follow music notation. In this question, the aural stimulus was divided into three sections—the first and last were sung aloud and the second was silent. Students were required to follow the notation and indicate which notes were silent.

Indices of Difficulty and Discrimination. The difficulty and discrimination indices for item 4 were .73 and .41, respectively. This difficulty index was greater than the value obtained in Field Study 1 (p = .06). The increase was attributed to the addition of a practise question of similar structure. In addition, the difficulty index obtained for item 4 in Field Study 3 was probably higher than that obtained for item 14 in Field Study 1 because the former melody composed in 6/8 meter was replaced by a melody in 4/4 meter (see Figure 45: frames 1 and 2). More specifically, since the tempo of the beat was held constant (quarter note = 76 beats/minute in 4/4 meter; dotted quarter note = 76 beats/minute in 6/8 meter) the duration of eighth notes in 6/8 meter was shorter, thus resulting in the increase in difficulty of item 14.



[a] \ indicates target notes.

Figure 45. Item 4: Comparison of Test Versions 3 and 4.

Item Revision. While the difficulty index obtained for item 4 meets the acceptable standard, it was felt that further revision was needed. Therefore, the melody employed in Test Version 3, composed of quarter notes and half notes (see Figure 45: frame 2), was revised so that it employed an eighth-note configuration (see Figure 45: frame 3). It was anticipated that it would be more difficult for students to identify the silent notes in the melody employed in Test Version 4 than the melody employed in Test Version 3.

Items 5 and 6 (3 and 4)

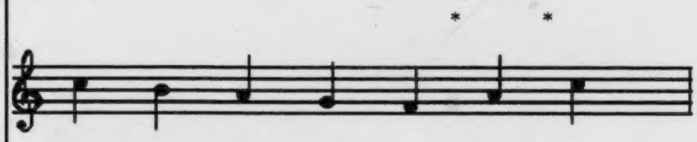
Items 5 and 6 were designed to assess a student's ability to identify all the notes which moved by a leap (item 5) and a step (item 6).

Indices of Difficulty and Discrimination. The difficulty indices for items 5 and 6 were .18 and .13, respectively. The corresponding point-biserial indices of discrimination were .25 for item 5 and .29 for item 6.

Item Revision. Despite revisions made to directions to students and melodic structure, the difficulty indices were below the suggested lower limit of .20. The low difficulty may be attributable to confusion in the terminology used to describe melodic movement. A review of the literature revealed inconsistencies in the terminology used to describe intervals larger than a step (ie; the intervals identified in item 5 using the term "leap"). For example, in the grade 1 teachers' guide of the series Musicanada, pitches were described as moving by step or skip (see Figure 46). In this sense, skips referred to intervals of a third. In the grade 3 edition of this series, skips were identified as intervals of a third, fourth, fifth, or sixth. In the grade 4 edition of the series Music and You, intervals between notes were identified as steps, skips, or repeats. Since leaps were not mentioned, it was assumed that skips referred to intervals greater than a second. Nye and Nye (1992) identified disjunct motion as notes that move by skip or leap. These authors did not clarify further the meaning of these terms. Similarly, Hackett and Lindeman (1988) defined skips as intervals larger than a step. They did not use the term leap. Conversely, Hoffer and Hoffer (1987) used the term leap to define intervals larger than a step. Hoffer and Hoffer's terminology was used in the current study. However, considering the discrepancies in terminology found in the literature, directions were modified to inform students that a leap is sometimes called a skip.

It is also possible that difficulty levels for both items were low because students did not understand how to circle the notes. Consequently, in Test Version 4, students were instructed to identify the pairs of notes that were separated by a leap (see Figure 47, frame

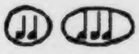
3) or by a step (see Figure 48, frame 2). Further, the diagrammatic representations of circled notes employed in Test Version 3 were deleted.



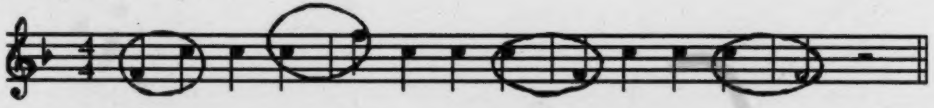
* indicates skipping notes.

Figure 46. Musicanada: Steps and skips.

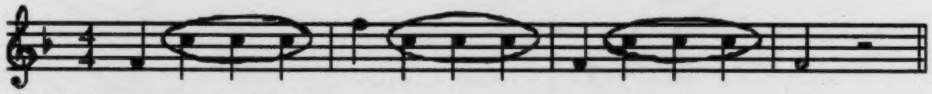
Test Version 3
a. Directions to Students

Listen to the melody once. Follow the music written below and circle all the notes  that move by a **leap**.

b. Expected Response



c. Incorrect Response Set



Test Version 4: Revised Form

Look at the music shown below and circle all the pairs of notes that are separated by a **leap**. A leap is sometimes called a skip.

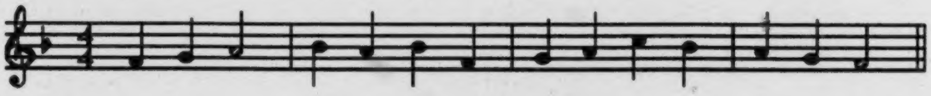
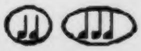


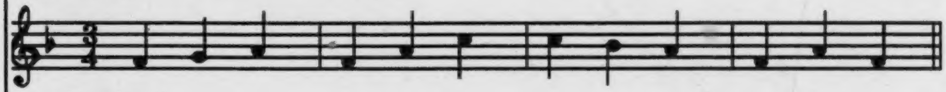
Figure 47. Item 5: Comparison of Test Versions 3 and 4.

In item 5 students were required to circle the notes that were separated by a leap. As illustrated in Figure 47 (frame 2), 27 students circled all the notes that repeated. This response suggested that either students differentiated between repeating notes and leaping notes or they confused the two concepts. Consequently, for purposes of scoring and interpretation, the melody employed in Test Version 3 was replaced by a melody with no repeated notes (see Figure 47: frame 3). It was also necessary to change the melody employed in item 6 so as to accurately reflect the specifications of the revised directions (see Figure 48).

Test Version 3

Listen to the melody once. Follow the music written below and circle all the notes

 that move by a **step**.



Test Version 4

Look at the music shown below and circle all the pairs of notes that are separated by a **step**.

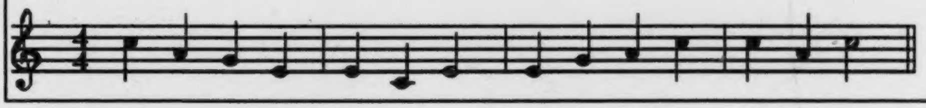


Figure 48. Item 6: Comparison of Test Versions 3 and 4.

Items 13 and 14 (25 and 26)

Items 13 and 14 were designed to assess a student's ability to notate an aurally-presented phrase.

Indices of Difficulty and Discrimination. The difficulty indices for items 13 and 14 were .12 and .21; point-biserial correlations were .29 and .31, respectively.

Item Revision. Items 13 and 14 were not revised. However, based on the low indices of difficulty obtained in Field Study 3, the requirements of these items were explained by means of a practise question.

Items 15 and 16 (15 and 16)

Items 15 and 16 were designed to assess a student's ability to identify the measure in which the aural presentation of a melody differed from the notational representation.

Indices of Difficulty and Discrimination. Items 15 and 16 were presented as four-option multiple-choice items. The difficulty indices obtained for the correct options were .60 and .78, respectively. The corresponding point-biserial correlations were .46 and .41.

If it can be expected that the students who did not select the correct answer were approximately evenly distributed across the three incorrect options, then, for item 15, the distribution of p-values for the incorrect options are acceptable when expected fluctuations due to random error are considered. Likewise, the point-biserials are acceptable given the theoretical upper limit as determined by the obtained p-value (.60).

For item 16, the second incorrect option appeared to attract a greater portion of students than the remaining two incorrect options (.12 versus .06 and .04). The discrimination indices were acceptable taking into account the upper theoretical limits.

Item Revision. Considering the easiness of item 16 relative to item 15, it is possible that students more easily identified the major tonic triad employed in the first two measures of the melody employed in item 16 than the melodic motive in the first measure of item 15. As illustrated in Figure 49, this portion of the melody employed in item 16 was replaced with less familiar motives—the major tonic triad with a passing tone on the second degree of the scale (measure one) and an open fifth interrupted with tones on the second and fourth degrees of the scale (measure two). As well, the melodic progression in the fourth measure of item 16, comprised of quarter notes and half notes in Test Version 3, was replaced with melodic progressions incorporating eighth notes (see Figure 49). This was done in an attempt to increase the likelihood that students would choose this option. Wanting to maintain the comparability of items 15 and 16, the same change was made to item 15 (see Figure 50).

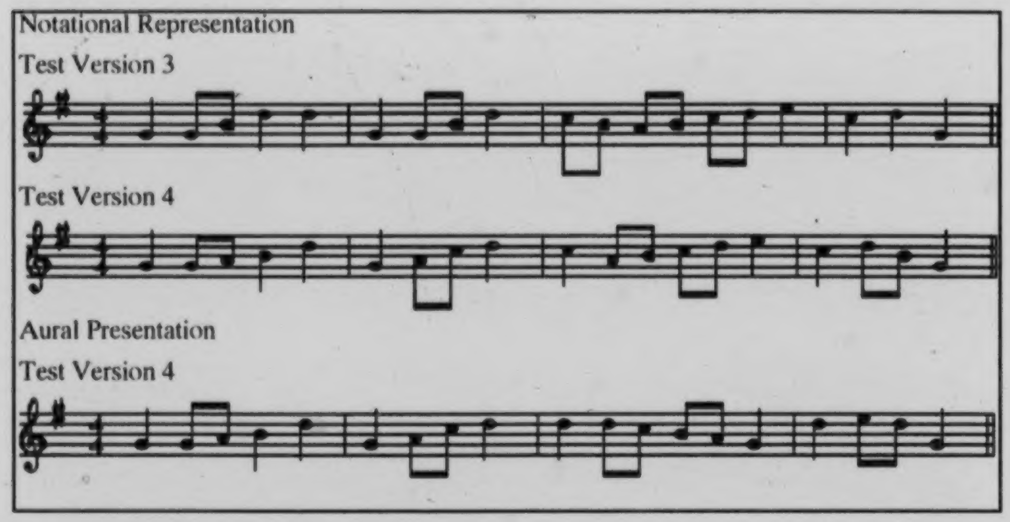

Notational Representation
Test Version 3
Test Version 4
Aural Presentation
Test Version 4

Figure 49. Item 16: Aural and visual representation.

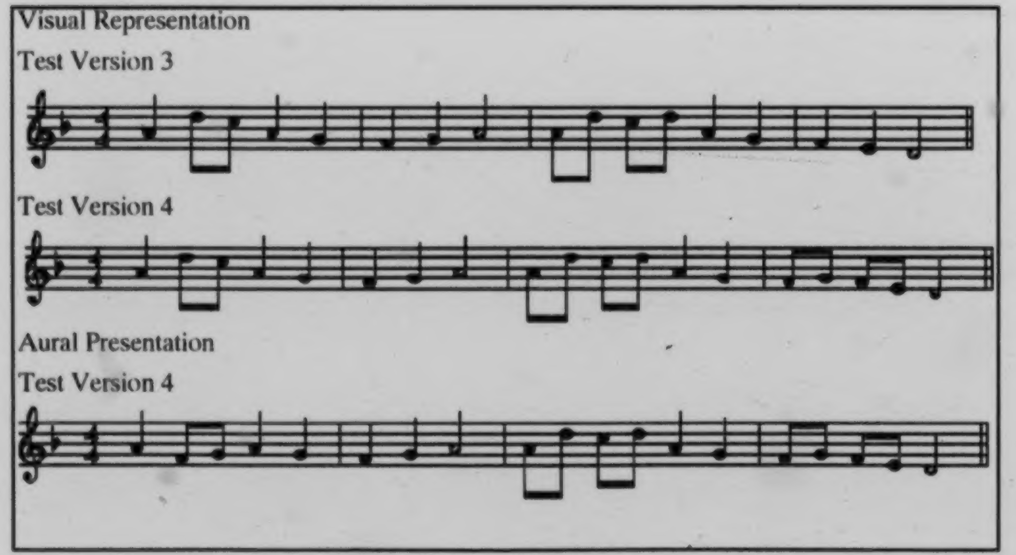

Visual Representation
Test Version 3
Test Version 4
Aural Presentation
Test Version 4

Figure 50. Item 15: Aural and visual representation.

Items 17 and 18 (1 and 2)

Items 17 and 18 were designed to assess a student's ability to identify the contour of a given melody.

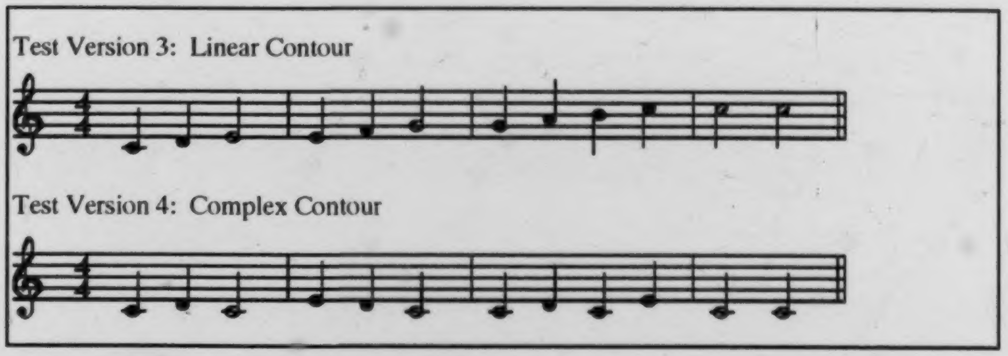
Indices of Difficulty and Discrimination. The difficulty indices for items 17 and 18 were .94 and .61, respectively. The low index of discrimination obtained for item 16, $r_{pbis} = .18$, was attributed to the high index of difficulty obtained for the item. The point-biserial index obtained for item 18 was .41.

It is interesting to note that p-levels obtained for items in which students were presented with melodies written in traditional music notation (Field Study 3) were higher than p-levels obtained for items in which students were presented with melodies written in diagrammatic form (Field Study 1: p = .63 and .65 for items 1 and 2, respectively). It is possible that the students were more accustomed to associating music heard aurally with notes on a staff than with other forms of diagrammatic representation. Thus, it would seem that the revised items employed in Field Study 3 are more appropriate for a test which assesses a student's ability to read music notation. It must be noted, however, that student response may also have been influenced by changes to directions implemented in Field Study 2.

Item Revision. Correct options and distractors were based on Long (1977) who described overall contour in terms of M-shaped and V-shaped melodies. The third option, a linear melody, was developed as a third alternative for use in the current study. The incorrect option which contained the linear melody was essentially non-functioning, particularly for item 18. In this case, while it was expected that about 20% of the students would select option A, only 8% did. In consideration of this, it seems likely that the linear representation is not suitable for multiple-choice items at the grade 5 level. Consequently, the linear option for each item was replaced by a melody with several changes in contour (denoted as complex contour) (see Figure 51 and Figure 52).

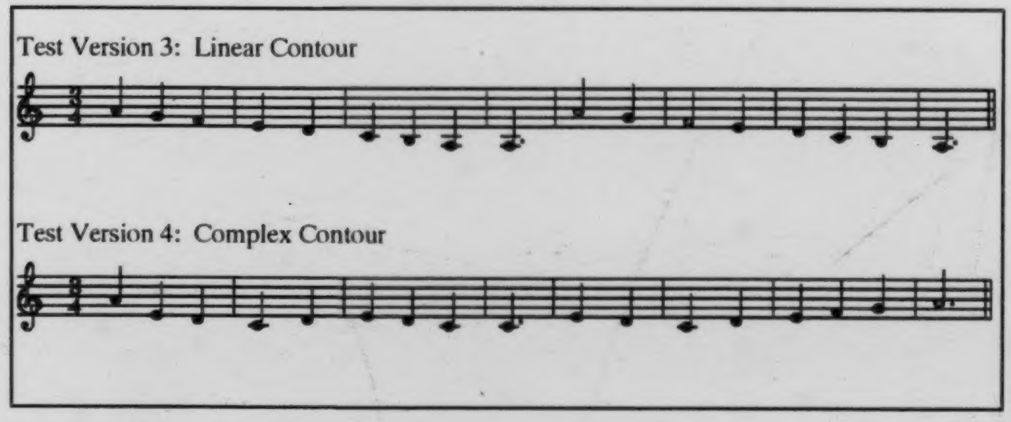
In item 17, the correct option contained a V-shaped melody. Conversely, in item 18, the correct option was an M-shaped melody. The difference in difficulty levels corresponds to Boltz and Jones' (1986) observation that melodies with several contour changes were more difficult to discern than melodies with fewer changes in contour.

Further, in the case of item 18, 31% of the students chose the V-shaped option, albeit incorrectly, lending further support to the observation of Boltz and Jones (1986).


Test Version 3: Linear Contour

Test Version 4: Complex Contour

Figure 51. Item 17: Linear and complex contour.


Test Version 3: Linear Contour

Test Version 4: Complex Contour

Figure 52. Item 18: Linear and complex contour.

Items 19 and 20 (17 and 18)

Items 19 and 20 were designed to assess a student's ability to identify the notational equivalent of an aural melody.

Indices of Difficulty and Discrimination. The difficulty indices obtained for items 19 and 20 were .83 and .42, respectively. Point-biserial correlations for the correct response, .37 and .26 respectively, were within the accepted guideline of .20.

The internal structure of item 19 was examined in order to determine why such a large proportion of students answered this item correctly. Option B, chosen by 13% of the students, obtained an acceptable level of discrimination ($r_{pbis} = -.40$). The low discrimination of option C ($r_{pbis} = -.03$) was likely due to the low proportion of students who chose this option (4%).

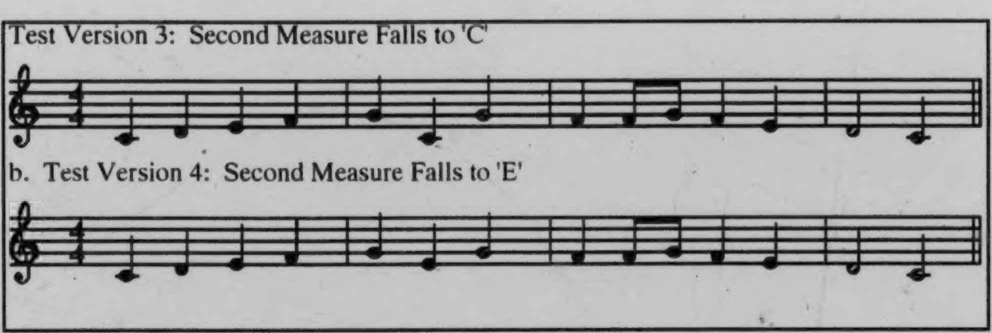
The greater difficulty of item 20, as compared to item 19, was attributed to the melodic structure of the incorrect options. Option B, an incorrect option in which the contour varied from that of the correct response, was chosen by 16% of the students and discriminated at $r_{pbis} = -.10$. Considering that the first two measures of option B differed from the correct response in terms of contour, it was expected that this option would discriminate between students to a greater extent. Option A, an incorrect option, was chosen by 43% of the students. This was similar to the correct response, option C, which was chosen by 42% of the students. Option A discriminated negatively ($r_{pbis} = -.19$) and option C discriminated positively ($r_{pbis} = .26$). Both options had the same contour. Therefore, students were required to discriminate exact interval size in order to correctly equate the aural stimulus to its' notational equivalent—a task performed more accurately by the high-scoring students.

Item Revision. For item 19, it is possible that students dismissed option C because, in comparison with the correct response, the transformation in contour included a note on a ledger line (middle 'C' in measure 2). Accordingly, the item was revised so that the melody falls to 'E' in measure 2 (see Figure 53). In terms of interval, the differences between corresponding notes in option A, the correct response, and option B, an incorrect alternative, may have been too wide. Therefore, interval differences of fifths, sixths, and sevenths were reduced to intervals of a third (see Figure 54).

For item 20, the nearly equal proportions of students who chose option A or option C indicated that the notational form of these melodies did not differ sufficiently. In Test Version 3, option A differed from option C on one note only (see Figure 55, frame 2). In

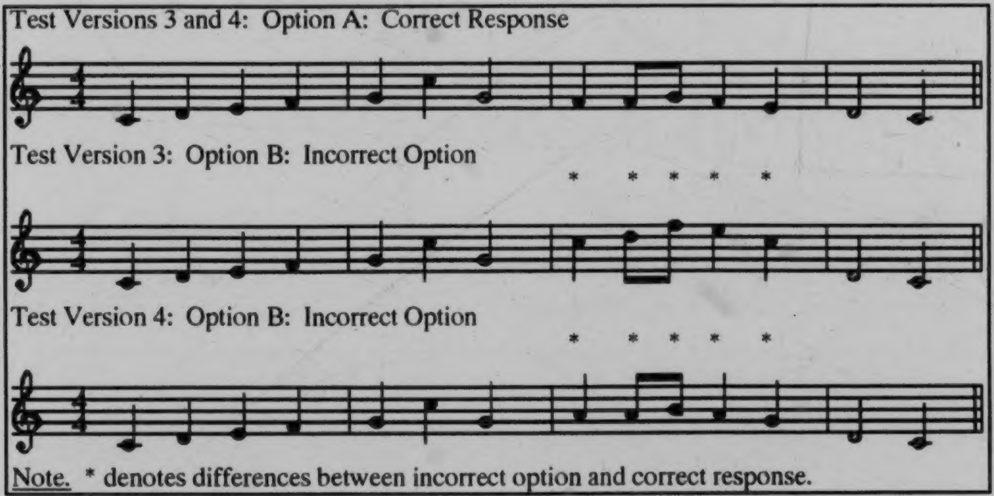
Test Version 4, option A differed from option C on two notes (see Figure 55, frame 3). The difference between the aural stimulus and the notational representation was accentuated by the high 'E' in the notational example (the second of two notes in which the notation differed from the melody heard by the students).

It appeared that both items 19 and 20 were susceptible to test-wiseness. Students dismissed the option in which the contour differed from the other two options in measure two and chose their response from the remaining two options which differed notationally in measure three. An additional option which differed from the correct response in measure three was included in each item (see Figures 56 and 57).


Test Version 3: Second Measure Falls to 'C'

b. Test Version 4: Second Measure Falls to 'E'

Figure 53. Item 19: Comparison of option C in Test Versions 3 and 4.

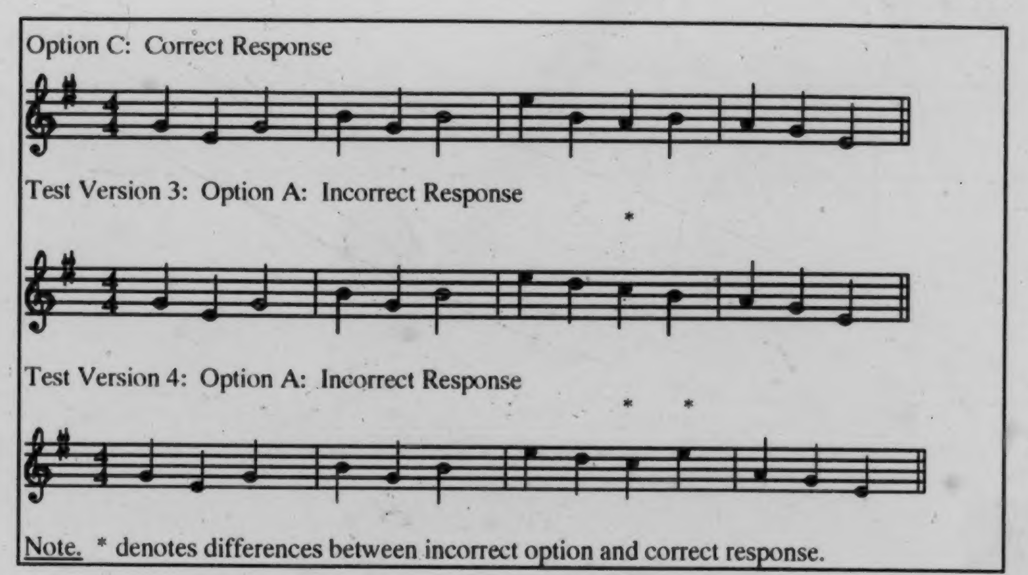

Test Versions 3 and 4: Option A: Correct Response

Test Version 3: Option B: Incorrect Option

Test Version 4: Option B: Incorrect Option

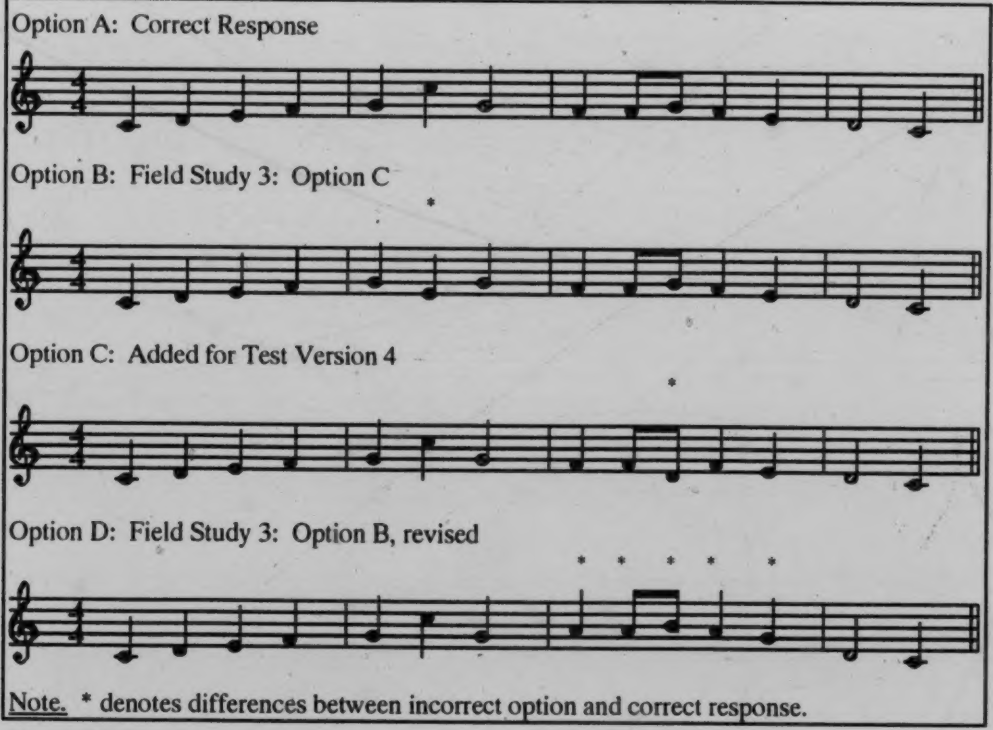
Note. * denotes differences between incorrect option and correct response.

Figure 54. Item 19: Comparison of option B in Test Versions 3 and 4.


Option C: Correct Response
Test Version 3: Option A: Incorrect Response
*
Test Version 4: Option A: Incorrect Response
* *

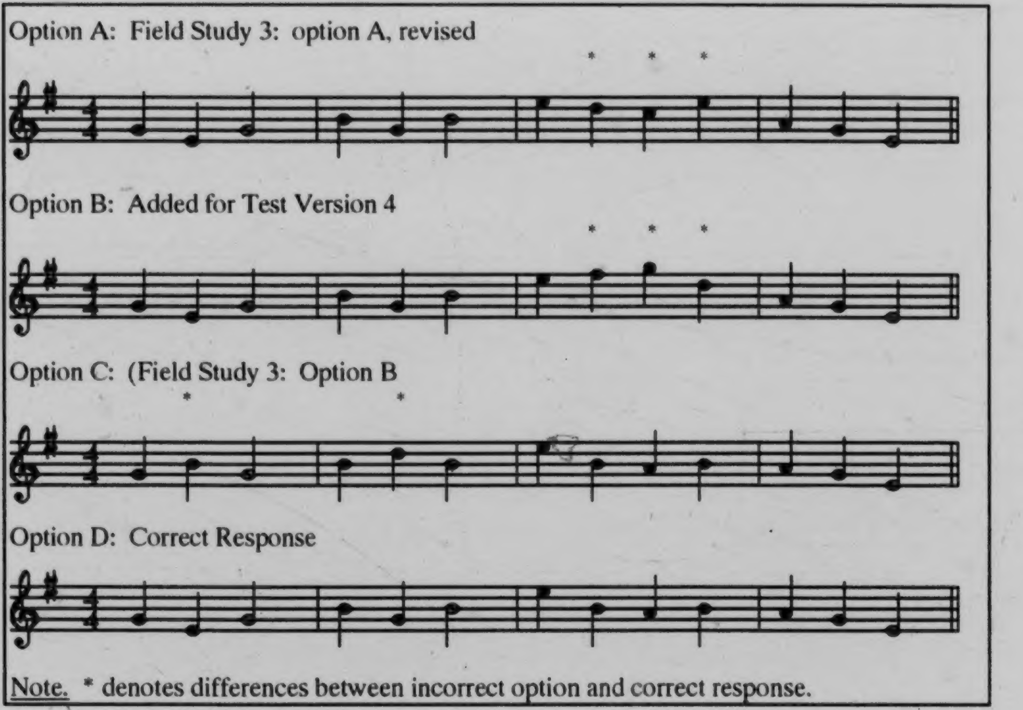
Note. * denotes differences between incorrect option and correct response.

Figure 55. Item 20: Comparison of option A in Test Versions 3 and 4.


Option A: Correct Response
Option B: Field Study 3: Option C
*
Option C: Added for Test Version 4
*
Option D: Field Study 3: Option B, revised
* * * * *

Note. * denotes differences between incorrect option and correct response.

Figure 56. Item 19: Test Version 4


Option A: Field Study 3: option A, revised
Option B: Added for Test Version 4
Option C: (Field Study 3: Option B
Option D: Correct Response

Note. * denotes differences between incorrect option and correct response.

Figure 57. Item 20: Test Version 4

Items 23 and 24 (21 and 22)

Items 23 and 24 were designed to assess a student's ability to identify a repeating melodic sequence.

Indices of Difficulty and Discrimination. The difficulty indices obtained for items 23 and 24 were .45 and .55, respectively. The point-biserial index of discrimination obtained for item 23, .11, was below the accepted guideline of .20. The index of discrimination obtained for item 24 was .39.

Point-biserial correlations were examined in order to assess the internal functioning of the items. In item 23, the correct response (option C) discriminated positively. However, at $r_{pbis} = .11$, this coefficient was below the desired level. Option A, chosen by 20% of students, did not discriminate between the overall performance of those students who chose that option and those students who did not choose that option ($r_{pbis} = .00$).

This seemed surprising considering that this option was dismissed easily by students in Field Study 2 because the melody constantly descended in the last two measures. Option B discriminated negatively (r_{pbis} = -.09). Contrary to expectations, this option, which differed from the correct response in terms of both contour and interval, was chosen by 35% of the students. In item 24, for option A (the correct response), the point-biserial index of discrimination was .39. The point-biserial indices of discrimination for options B (chosen by 13% of students) and C (chosen by 30% of students) were -.15 and -.26, respectively.

Item Revision. Considering the possibility that students did not understand how to answer items 23 and 24, the instructions were revised again. For Field Study 4, this change is illustrated for item 23 in Figure 58. As shown, the intent of the revised directions was to explain more specifically how the notated melodies were constructed. Item 24 was revised similarly.

Test Version 3

Listen to each melody once. The first measure of each melody below begins with the same melodic **pattern** or **sequence**. Measures 2 and 3 are different. Circle the letter above the **one** melody which **best continues** the repeating melody.

Test Version 4

Look at the music shown below. The first measure of melodies **A, B,** and **C** all begin with the **same melodic pattern**. After that, each melody is different. Circle the letter above the **one** melody where the melodic pattern shown in the first measure is repeated in different places on the staff.

Figure 58. Item 23: Comparison of instructions to students in Test Versions 3 and 4.

Summary Description

Summary Test Statistics

The item scores were aggregated for each individual to produce a total score out of 24. For the sample of 77 students, the mean was 12.20 (50.8%), the median was 12.0 (50.0%), and the mode was 12.5. The scores ranged from 3 to 21, with a standard deviation of 4.7 (5.1%). The internal consistency, coefficient alpha (Cronbach, 1951), was .74; the corresponding SEM was 2.05. Taken together, the test as a whole was of medium difficulty. The scores were essentially symmetrically distributed around the center point. The internal consistency, while less than .80, was appropriate for this type of measure (Anastasi, 1988; Ebel & Frisbie, 1991).

Chapter Summary

In Field Study 3, Test Version 3 was administered to 77 students in four grade 6 classes in the Fall of 1994. Based on the data obtained in Field Study 3, items 3, 7, 8, 9, 10, 11, 12, 21, and 22 were not revised. Revisions to the remaining 13 items were described in the item-by-item analysis. Of these, directions to students were amended in items 5, 6, 23, and 24; melodies were modified in items 5, 6, 16, 17, 18, 19, and 20; the rhythmic material was revised in items 1, 4, 15, and 16; and the target note was adjusted in item 2. For items 13 and 14, a practise question was provided in an attempt to ensure that students understand how to proceed with these items. Revisions were incorporated in Test Version 4, a copy of which is provided in Appendix H.

However, another issue arose at this juncture, namely that items referenced to 5 tasks might be solved by attending solely to the visual component: the identification of steps or leaps (skips), the identification of solfa syllables, the identification of absolute note names, the specification of the home tone, and the recognition of sequential patterns. Consequently, the mode of presentation (aural/visual or visual only) employed for these items was not specified pending the results of Field Study 4.

CHAPTER 7

FIELD STUDIES 4 AND 5: MODE OF PRESENTATION

Overview

It was not possible to predict whether students would rely solely on the visual representation of the melody independent of its aural presentation. Therefore, due to the lack of reported research, all melodies were sung in the previous field studies. At this juncture it seemed prudent to examine the strategies employed by students when responding to the items that might be solved by attending solely to the visual presentation in the absence of an aural presentation. This issue was addressed in Field Study 4 using the procedures employed in Field Study 2 and selected items from Test Version 4.

Based on the findings from Field Study 4, changes in addition to those described in the previous chapter were made to the set of 24 items. This resulted in the formulation of Test Version 5. Test Version 5 was arranged in two main sections: the first was comprised of those items which were presented both aurally and visually (items 1 to 16) and the second contained those items which, as based on the findings of Field Study 4, were presented in visual mode only (items 17 to 24). As detailed in the second main section of Chapter 7, Test Version 5 was administered to students in two grade 5 classes (see Field Study 5). The primary purpose of this study was to examine how students answered the visual-only items within a group situation. In this respect, difficulty indices obtained in Field Study 5 were compared with those obtained in Field Study 3.

The chapter concludes with a summary of item revisions based on the data collected in Field Study 5.

Field Study 4: Protocol Analysis

Introduction

Analysis of the 12 clusters of items revealed that the items in 5 clusters might be solved by attending solely to the visual component. The five clusters are: identification of steps or leaps (skips), identification of solfa syllables, identification of absolute note names, specification of the home tone, and recognition of sequential patterns. The item from each of these five clusters that was administered in Field Study 2 was selected for administration in Field Study 4. This allowed a comparison between the item administered in visual-aural manner and the item administered using only the visual mode.

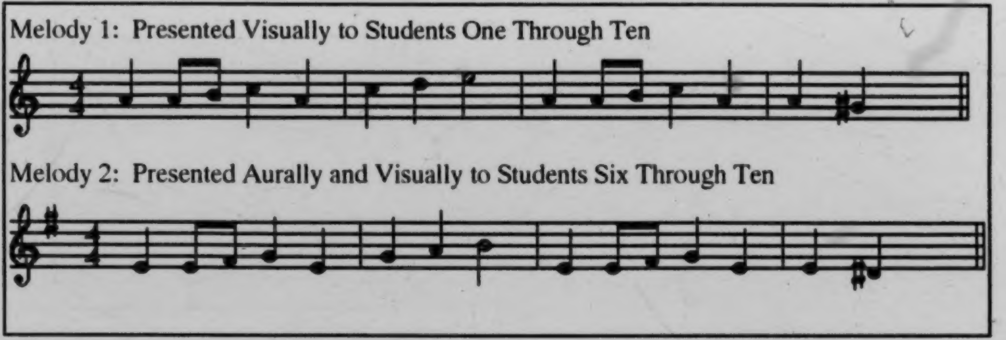
The items designed to assess the identification of solfa syllables, note names, and the home tone were not changed between Field Study 2 and Field Study 4. However, as described in the previous chapter, the remaining two items were revised. More specifically, the item in which students identified a repeated melodic pattern was clarified in terms of directions. In addition, the melody employed in the third option was revised in order to avoid the use of a ledger line note below the staff. The item in which students identified notes which were separated by a step was revised in terms of both directions to students and melodic composition. A copy of this short test, called Test Version 4S, is presented in Appendix H.

In January and February of 1995, Test Version 4S was administered to 10 grade 5 students, 5 in each of 2 schools located within School District 3. Each of the 2 teachers whose students participated in this study was requested to choose 5 students using the following criteria: 4 students within a range of average music achievement, and 1 student with above average achievement in music. Each student involved in this study returned a form signed by a parent or guardian indicating that the student was allowed to take part in this study. Using the same data-gathering procedures employed in Field Study 2, the 10 students were interviewed separately.

It was possible that strategies employed by students in the absence of an aural stimulus would differ from strategies employed when both components were present. Therefore, student responses to the items included in Test Version 4S were analyzed prior to making any comparisons with the results from Field Study 2 for these items.

Results

A detailed account of the data obtained in Field Study 4 is presented in Appendix I. In summary, the analysis of data revealed that the strategies employed by students when solving the items administered in the absence of the aural presentation of the melody differed little from strategies employed by students when solving these items in the presence of both an aural and a visual stimulus with one exception—the item in which students were required to identify the home tone of a given melody. With respect to this item (see Appendix I, item 10), two of the first five students interviewed in Field Study 4 referred to the sound of the melody or sang the melody aloud even though they were not given an aural stimulus. Consequently, to ascertain whether the inclusion of an aural stimulus would assist the students on this item, during the five remaining interviews students were presented with a second item in which the same melody used for the original item (written in the key of 'a' minor) was presented to students both aurally and visually in the key of 'e' minor (see Figure 59). Two of five students who were unable to answer the item correctly when presented with only the music notation were able to answer the item correctly when they were able to both see and hear the item. Therefore, the combined visual-aural presentation was retained for the pair of items designed to assess a student's ability to identify the home tone; the remaining four pairs of items were administered in visual mode only.


Melody 1: Presented Visually to Students One Through Ten

Melody 2: Presented Aurally and Visually to Students Six Through Ten

Figure 59. Home tone: Melodies presented in visual and aural/visual mode.

Field Study 5: Group Administration

Overview

As demonstrated in Field Study 4, students answered items referenced to four clusters by attending solely to the visual representation of the melody. To accommodate this change in presentation, Test Version 5 (see Appendix J) was arranged in two main sections. The first section contained the 16 items which were presented both aurally and visually and the second section included the 8 items which were presented visually.

Test Version 5 was administered to students in 2 grade 5 classes in March, 1995. All students attended one of two schools within School District 3 and each class was taught by one of two music specialists. Six students returned parental permission forms in School 1; 17 students returned parental permission forms in School 2. Therefore, due to the small number of students in School 1 who participated in this study, only 23 students completed the test.

Mode of Presentation

Items 17 to 24 were administered in aural-visual mode in Field Study 3 and visual mode only in Field Study 5. Further, as pointed out earlier, the items were moved so that all the visually-administered items were administered last. For example, items 5 to 10 in Test Version 3 were moved to become items 17 to 22 in Test Version 5. Items 23 and 24 were in the same position in both versions.

Assessment of Visual Mode Only

Presented in Table 19 are the item difficulty and discrimination indices for the eight items administered in visual mode only. These values are accompanied by their corresponding values from Field Study 3 in which an aural-visual presentation was used.

Examination of the pair of difficulty indices for each item in Table 19 reveals that, sampling error notwithstanding, the performance was higher for the visual mode for both items referenced to absolute note names, the performance was higher for only one of the

Table 19

Field Studies 5 (FS 5) - Visual Only and 3 (FS 3) - Visual-Aural: Indices of Difficulty (p) and Discrimination (r_{pbis}) for Items 17 - 24

				Options							
				A		B		C		D	
Item	Task	Format	Index	FS 5	FS 3	FS 5	FS 3	FS 5	FS 3	FS 5	FS 3
17 (5)	leap	WR	p	.22*	.18*	.78	.82				
			r_{pbis}	.26	.25	-.26	-.25				
18 (6)	step	WR	p	.13*	.13*	.87	.87				
			r_{pbis}	.40	.29	-.40	-.29				
19 (7)	names	WR	p	.61*	.40*	.39	.60				
			r_{pbis}	.47	.55	-.47	-.55				
20 (8)	names	WR	p	.83*	.65*	.17	.35				
			r_{pbis}	.47	.63	-.47	-.63				
21 (9)	solfa	WR	p	.57*	.18*	.43	.82				
			r_{pbis}	.33	.47	-.33	-.47				
22 (10)	solfa	WR	p	.39*	.45*	.61	.55				
			r_{pbis}	.28	.59	-.28	-.59				
23 (23)	sequence	MC	p	.22	.18	.30	.35	.48*	.45*		
			r_{pbis}	-.15	.00	-.15	-.09	.26	.11		
24 (24)	sequence	MC	p	.39*	.55*	.30	.13	.30	.30		
			r_{pbis}	.30	.39	-.26	-.15	-.06	-.26		

Note. * denotes correct response. MC = multiple-choice; WR = written response. Parentheses indicate items in Field Study 3.

items referenced to steps or leaps, one of two items referenced to solfa syllables, and one of two items referenced to sequential patterns. In two of these items, the identification of a step or a leap and the identification of sequential patterns, there was no difference in performance between the two modes of administration. Thus, it appeared that there was a slight advantage to using the visual mode only for the eight items. Consequently, in Test Version 6, these items (Test Version 5: items 17 to 24; Test Version 6: items 15 to 22) were administered in visual mode only.

Item Revision

The indices of difficulty and discrimination obtained in Field Study 5 are reported in Table 20. These indices were analyzed with particular emphasis on implications for item revision. Of the 24 items administered in Field Study 5, 11 items were not revised (items 3, 7, 8, 11, 12, 13, 14, 19, 20, 21, and 22). Changes were made to items 1, 2, 4, 9, 10, 15, 16, 17, 18, 23, and 24. Items 5 and 6 were deleted from the study.

The item-by-item analysis which follows is limited to a description of those items that were revised or deleted in consideration of the data obtained in Field Study 5. In this respect, results are discussed below for each cluster or pair of items. The exception is test item 4, which is discussed in the absence of item 3 (an item which was not revised). Unless specified otherwise, item numbers refer to the order in which items were administered in Test Version 5. Numbers in parentheses refer to the order in which the items were administered in Field Study 3. The analysis of the items not revised in Field Study 5 is provided in Appendix K.

Items 1 and 2 (Items 1 and 2 in Field Study 3)

Items 1 and 2 were designed to assess a student's ability to follow music notation. In these questions, only part of the melody was sung aloud. Students were required to draw a line through the final note they heard.

Table 20

Field Study 5: Indices of Difficulty and Discrimination

Item	Task	Format	Index	Options A	B	C	D	Revise
1 (1)	follow notes	WR	p	.83*	.17			x
			r_{pbis}	.16	-.16			
2 (2)	follow notes	WR	p	.83*	.17			x
			r_{pbis}	.47	-.47			
3 (3)	follow notes	WR	p	.61*	.39			
			r_{pbis}	.52	-.52			
4 (4)	follow notes	WR	p	.74*	.26			x
			r_{pbis}	.36	-.36			
5 (13)	dictation	WR	p	.22*	.78			x
			r_{pbis}	.35	-.35			
6 (14)	dictation	WR	p	.09*	.91			x
			r_{pbis}	.35	-.35			
7 11)	home tone	WR	p	.35*	.65			
			r_{pbis}	.29	-.29			
8 (12)	home tone	WR	p	.65*	.35			
			r_{pbis}	.28	-.28			
9 (15)	read notes	MC	p	.87*	.04	.09	.00	x
			r_{pbis}	.02	.05	-.06	.00	
10 (16)	read notes	MC	p	.09	.04	.78*	.09	x
			r_{pbis}	-.25	-.20	.31	-.06	
11 (17)	contour	MC	p	.09	.87*	.04		
			r_{pbis}	-.15	.25	-.20		
12 (18)	contour	MC	p	.22	.65*	.13		
			r_{pbis}	-.31	.28	-.02		
13 (19)	read notes	MC	p	.61*	.04	.30	.04	
			r_{pbis}	.47	.24	-.65	.12	

Note. * denotes correct response. MC = multiple-choice; WR = written response.

Numbers in parentheses refer to item order in Field Study 3.

Table 20

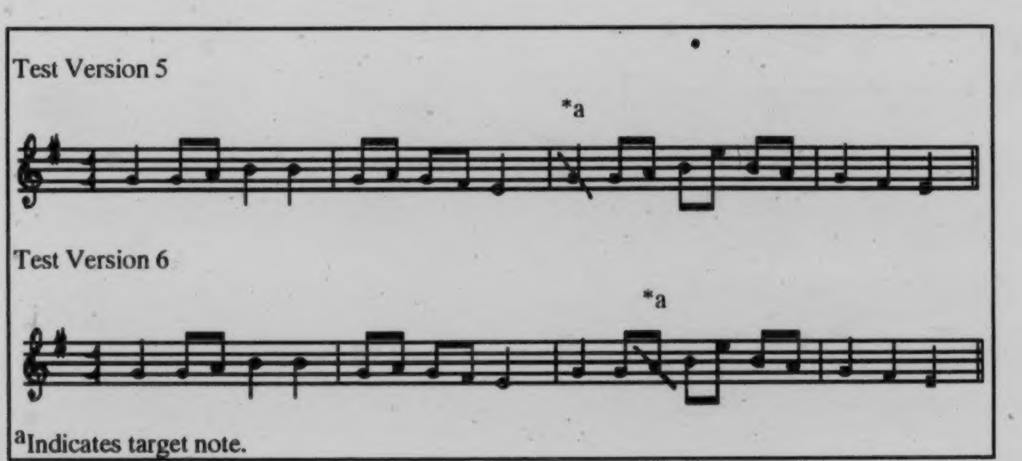
Field Study 5: Indices of Difficulty and Discrimination

				Options				
Item	Task	Format	Index	A	B	C	D	Revise
14 (20)	read notes	MC	p	.26	.04	.13	.57*	
			r_{pbis}	-.15	-.07	.33	-.06	
15 (21)	sharps	MC	p	.39	.09	.52*		x
			r_{pbis}	-.41	.12	.33		
16 (22)	flats	MC	p	.39*	.52	.09		x
			r_{pbis}	.52	-.42	-.15		
17 (5)	leap	WR	p	.22*	.78			x
			r_{pbis}	.26	-.26			
18 (6)	step	WR	p	.13*	.87			x
			r_{pbis}	.40	-.40			
19 (7)	names	WR	p	.61*	.39			
			r_{pbis}	.47	-.47			
20 (8)	names	WR	p	.83*	.17			
			r_{pbis}	.47	-.47			
21 (9)	solfa	WR	p	.57*	.43			
			r_{pbis}	.33	-.33			
22 (10)	solfa	WR	p	.39*	.61			
			r_{pbis}	.28	-.28			
23 (23)	sequence	MC	p	.22	.30	.48*		x
			r_{pbis}	-.15	-.15	.26		
24 (24)	sequence	MC	p	.39*	.30	.30		x
			r_{pbis}	.30	-.26	-.06		

Note. * denotes correct response. MC = multiple-choice; WR = written response. Numbers in parentheses refer to item order in Field Study 3.

Indices of Difficulty and Discrimination. The difficulty indices obtained for both items 1 and 2, .83, exceeded the suggested level of .80; point-biserial correlations were .16 and .47, respectively.

Item Revision. Comparison with the difficulty indices obtained in Field Study 3 (item 1, p = .84; item 2, p = .83) indicated that revisions meant to make the items more difficult did not have the intended effect. In Field Study 3, for item 1 the simple rhythmic structure of the melody was revised to include eighth notes; for item 2 the target note was placed on a musically less important beat. Considering the data obtained in Field Study 5, these items were revised once more. As presented in Field Study 5, in item 1 students were required to identify a note on the first beat of measure three. In the revised form, students were required to identify a note on a musically less important beat—the second half of the second beat in the third measure (see Figure 60). In item 2, the quarter note-eighth note structure heard in Test Version 5 was replaced by a series of eighth notes (see Figure 61). In Test Version 5, students were required to identify a note on the fourth beat of the second full measure of a melody written in 6/8 time. In Field Study 6, students were required to identify a note on a musically less important beat—in 6/8 time, the fifth beat in the second full measure.



[a]Indicates target note.

Figure 60. Item 1: Comparison of target notes in Test Versions 5 and 6.

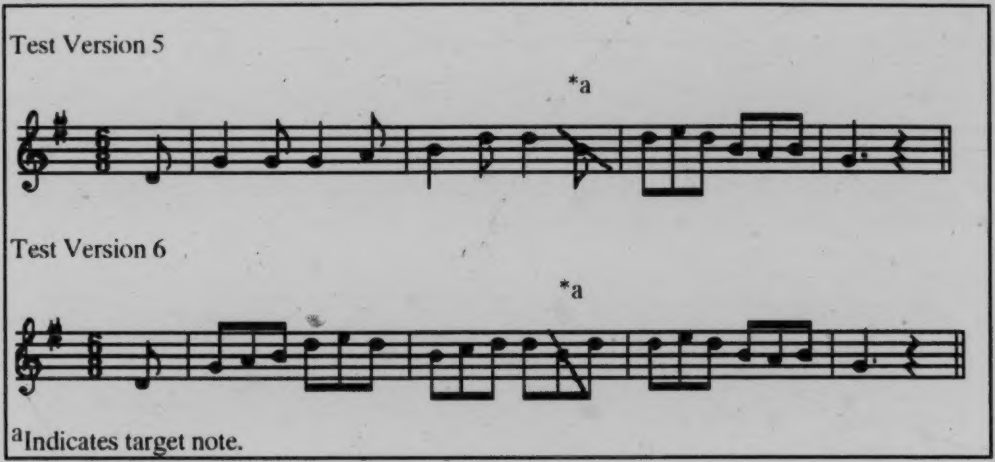


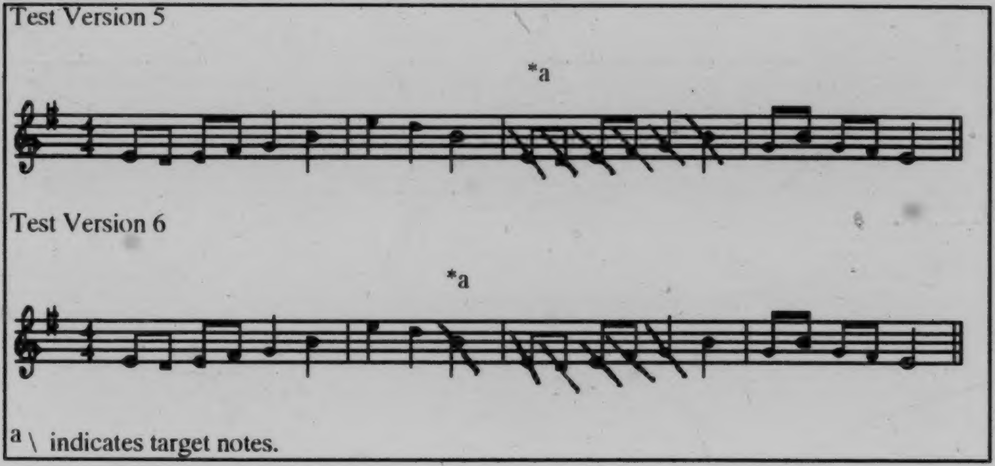
Figure 61. Item 2: Comparison of target notes in Test Versions 5 and 6.

Item 4 (Item 4 in Field Study 3)

Item 4 was designed to assess a student's ability to follow music notation. In this question, the aural stimulus was divided into three sections—the first and last were sung aloud and the second was silent. Students were required to follow the notation and indicate which notes were silent.

Indices of Difficulty and Discrimination. The difficulty index obtained for item 4, .74, approached the upper limit of the accepted guideline of .80. The point-biserial index of discrimination was .36.

Item Revision. In Field Study 3, the melody employed for item 4, composed of quarter notes, was revised so that it employed an eighth-note configuration. This revision did not make the item more difficult (Field Study 3: $p = .73$). Therefore, in a further attempt to make the item more difficult, the number of silent beats was increased from 4 to 5 (see Figure 62).


Test Version 5

Test Version 6

[a] \ indicates target notes.

Figure 62. Item 4. Target notes.

Items 5 and 6 (Items 13 and 14 in Field Study 3)

Items 5 and 6 were designed to assess a student's ability to notate an aurally-presented phrase.

Indices of Difficulty and Discrimination. The difficulty index for item 5, .22, approached the lower level of the accepted guideline of .20. The difficulty index for item 6, .09, was below this accepted guideline. The point-biserial indices of discrimination, .35 for both items, exceeded the accepted guideline of .20.

Data analyses conducted to this point in the study consistently indicated that these items were too difficult. An examination of the responses obtained in Field Studies 3 and 5 revealed that students perceived the overall contour of the melodies but were unable to discern exact interval. In Field Study 3, 18 students who incorrectly responded to item 5 (item 13 in Test Version 3) and 32 students who answered item 6 incorrectly (item 14 in Test Version 3) were able to notate a melody in which some semblance of the contour was retained. Likewise, in Field Study 5, 6 students who incorrectly responded to item 5 and 14 students who responded incorrectly to item 6 were able to notate a melody which resembled the melody they heard in terms of contour. Beyond this, in all probability little additional information could be gained by re-administering these items in their current form.

Item Revision. Due to low indices of difficulty, these items were not included in the revised version of the test.

Items 9 and 10 (Items 15 and 16 in Field Study 3)

Items 9 and 10 were designed to assess a student's ability to identify the measure in which the aural presentation of a melody differed from the notational representation.

Indices of Difficulty and Discrimination. The difficulty indices for items 9 and 10 were .87 and .78, respectively. The low index of discrimination obtained for item 9, $r_{pbis} = .02$, was less than the desired level, .20. The point-biserial index of discrimination for item 10 was .31.

Item Revision. Despite the revisions made to items 9 and 10, the items were still too easy. Seemingly the students had no problems identifying the correct response. Consequently, for both items 9 and 10, the notation for the correct option was revised so that it more closely resembled the aural presentation (see Figures 63 and 64).

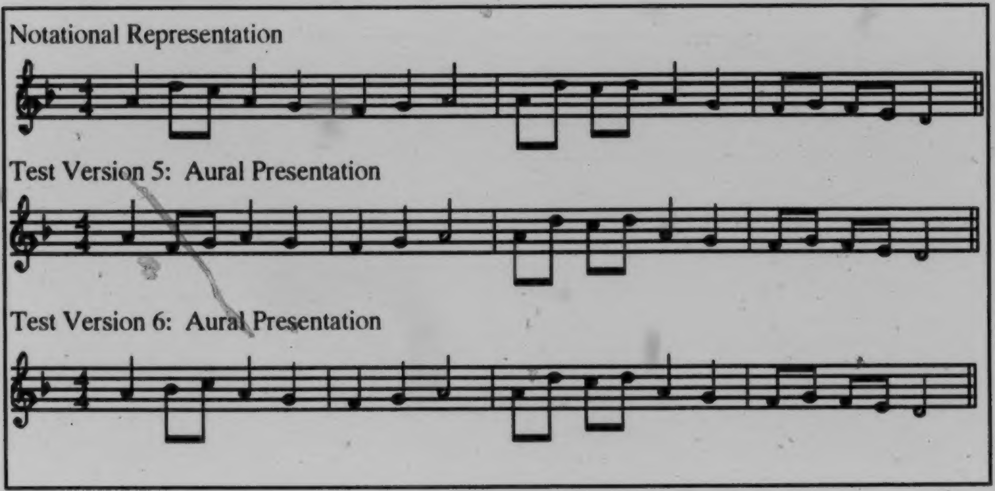

Notational Representation
Test Version 5: Aural Presentation
Test Version 6: Aural Presentation

Figure 63. Item 9: Comparison of aural representation in Test Versions 5 and 6.

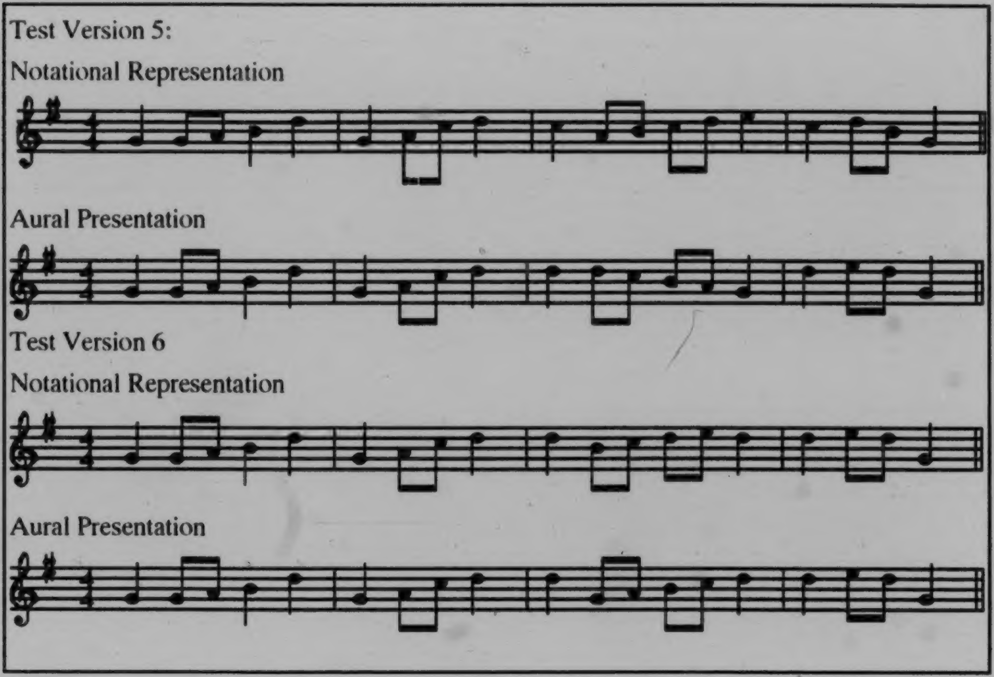

Test Version 5:
Notational Representation
Aural Presentation
Test Version 6
Notational Representation
Aural Presentation

Figure 64. Item 10: Comparison of Test Versions 5 and 6.

Items 15 and 16 (Items 21 and 22 in Field Study 3)

Items 15 and 16 were designed to assess a student's ability to identify the sharps or flats in an aural presentation of a melody.

Indices of Difficulty and Discrimination. The difficulty indices for items 15 and 16 were .52 and .39, respectively. Point-biserial indices of discrimination were .33 and .52.

In item 15, students were required to identify a melody which had been altered by sharps. Option A, in which the melody was altered by flats, had a discrimination index of $r_{pbis} = -.41$ and was chosen by 39% of the students. The melody with no sharps or flats (option B) discriminated positively ($r_{pbis} = .12$) and was chosen by 9% of the students.

In item 16, students were required to identify a melody which had been altered by flats. Option B, which contained sharps, had a discrimination index of $r_{pbis} = -.42$ and option C, which contained no sharps or flats, had a discrimination index of $r_{pbis} = -.15$. It is interesting to note that 9% of the students chose this option—the same number as the option with a parallel function in item 15 (option B). Considering that this option discriminated positively in item 15, it appears that these items were subject to test-wiseness (a conclusion also indicated in the data obtained for item 16 in Field Study 3). Due in part to practise effects, by the time the second item was presented high scoring students had determined that they were expected to differentiate between sharps and flats. More specifically, some of the high scoring students who chose the option with no sharps or flats in item 15 (option B: $r_{pbis} = .12$) chose between the options with either sharps or flats in item 16. Low scoring students continued to choose among all three options—thus contributing to the negative discrimination ($r_{pbis} = -.15$) of option C in the latter item.

Item Revision. To account for this possible problem of test-wiseness susceptibility, instructions to listen for the sharps or flats were deleted in favor of the directions employed in items 11 through 14. As such, in Test Version 6, students were instructed to circle the letter above the one melody which is the same as the melody they hear.

Items 17 and 18 (Items 5 and 6 in Field Study 3)

Items 17 and 18 were designed to assess a student's ability to identify all the notes which moved by a leap (item 17) or a step (item 18).

Indices of Difficulty and Discrimination. The difficulty index obtained for item 17, .22, approached the lower limit of the accepted guideline of .20. The difficulty index obtained for item 18, .13, was below this accepted guideline. The point-biserial indices of discrimination, .26 for item 17 and .40 for item 18, exceeded the accepted guideline of .20.

Item Revision. Revisions were made to make the items easier. In Field Study 5, the students were instructed to circle all the pairs of notes that were separated by a leap. It is possible that students did not know what was meant by the phrase "pairs of notes". Therefore, the instructions for students were revised so that they were requested to circle all the places in the music where two notes beside each other are separated by a leap (item 17) or a step (item 18).

Items 23 and 24 (Items 23 and 24 in Field Study 5)

Items 23 and 24 were designed to assess a student's ability to identify a repeating melodic sequence.

Indices of Difficulty and Discrimination. The difficulty indices for items 23 and 24 were .48 and .39. Both items discriminated positively (r_{pbis} = .26 and .30, respectively).

An internal analysis of item 23 revealed that the incorrect alternatives were functioning properly (option A: r_{pbis}= -.15; option B: r_{pbis} = -.15). For item 24, option B was functioning as expected (r_{pbis} = -.26). Option C was essentially nondiscriminating (r_{pbis} = -.06).

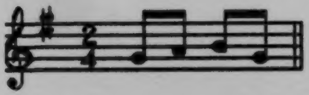
Item Revision. Considering that, for item 24, option C did not discriminate properly, the directions to students were altered in an attempt to clarify the specifications of this task (see Figure 65).

Test Version 5

Look at the music shown below. The first measure of melodies **A, B,** and **C** all begin with the **same melodic pattern**. After that, each melody is different. Circle the letter above the **one** melody where the melodic pattern shown in the first measure is repeated in different places on the staff.

Test Version 6

Look at the following melodic pattern.



The first measure of each melody begins with this pattern- after that, each melody is different. Circle the letter above the **one** melody where this melodic pattern is repeated in different places on the staff.

Figure 65. Item 23: Comparison of directions to students in Test Versions 5 and 6.

Chapter Summary

As described in Chapter 7, Field Studies 4 and 5 examined selected items in terms of mode of presentation. In Field Study 4, the strategies employed by students when responding to the items that, theoretically, could be answered by attending solely to the visual component in the absence of an aural presentation were examined by means of protocol data obtained through test administration on an individual basis. Analysis of data revealed that the strategies employed by students when solving 4 of the 5 items administered in the absence of the aural presentation of the melody (see Field Study 4) differed little from strategies employed by students when solving these items in the presence of both an aural and a visual stimulus (see Field Study 2). In the case of the fifth item, it was found that more students were able to identify the home tone correctly when the melody was presented both visually and aurally. Therefore, for this pair of items, the combined visual-aural presentation of the melody was retained.

In consideration of the findings of Field Study 4, a fifth test version was formulated—one in which selected items were presented in visual mode only. Tasks included in this group were: the identification of steps and leaps, the identification of solfa syllables, the identification of absolute note names, the specification of the home tone, and the recognition of sequential patterns.

Test Version 5 was arranged in two main sections: the first was comprised of those items which were presented both aurally and visually (items 1 to 16) and the second contained those items which, as based on the findings of Field Study 4, were presented in visual mode only (items 17 to 24). Test Version 5 was administered to two grade 5 classes in School District 3. A comparison of difficulty indices obtained in Field Study 5 and Field Study 3 revealed that there was a slight advantage to administering items 17 to 24 in visual mode only.

Based on the data obtained in Field Study 5, items 3, 7, 8, 11, 12, 13, 14, 19, 20, 21, and 22 were not revised. Items 5 and 6 were deleted from the study. Revisions were made to the remaining 11 items. Directions to students were modified in items 15, 16, 17, 18, 23, and 24; melodies were modified in items 9 and 10; the rhythmic material was revised in item 2; and the target notes were adjusted in items 1, 2, and 4. These revisions were incorporated in Test Version 6, a copy of which is provided in Appendix L.

Chapter 8

Field Study 6: Test Version 6 (Group Administration)

Overview

The purpose of Field Study 6 was to examine the functioning of test items assigned to Test Version 6 when administered to a sample of sufficient size so that indices of difficulty and discrimination could be estimated with relative stability. As well, data obtained in this administration were used to estimate the internal consistency of the items when viewed as an aggregate. Lastly, further item revision was guided by an analysis of indices of difficulty and discrimination obtained from the data collected in this study.

The presentation and interpretation of the data obtained in Field Study 6 is structured within this chapter in five main sections: first, a summary of the indices of difficulty and discrimination obtained in Field Study 6 with a focus on implications for item revision; second, a summary of the overall functioning of the items as revealed in summary test statistics; third, an examination of the students' opportunity to learn; fourth, a description of the sample of students involved in this study with particular emphasis on experiential background; and, fifth, an investigation of the correlations among items as demonstrated by means of factor analysis.

Description of the Sample

Test Version 6, which contained a total of 22 items, was administered to 214 grade 5 students in April and May, 1995.

The number of students exceeded the minimum of 200 suggested by Crocker and Algina (1986) when examining a near final form of a test. All students attended 1 of 7 schools within School District 1, 2, or 3 and all students were taught by a music specialist. Each student involved in this study returned a form signed by a parent or guardian indicating that the student was allowed to take part in the study.

Results

Item Revision

Indices of difficulty and discrimination obtained in Field Study 6 are reported in Table 21 for each of the 22 items. These indices were analyzed with particular emphasis on implications for item revision. Of the 22 items administered in Field Study 6, 17 items had satisfactory indices and, consequently were not revised (items 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 17, 18, 19, and 20). A summary analysis of the data obtained for these items is presented in Appendix M. The values obtained for the remaining items (11, 13, 16, 21, and 22) were not satisfactory. The item-by-item analysis which follows is limited to a description of data obtained for these five items. The item numbers refer to the order in which items were administered in Test Version 6.

Item 11

Item 11 was designed to assess a student's ability to identify the measure in which the aural presentation of a melody differed from its notational representation.

Indices of Difficulty and Discrimination. The difficulty index for item 11 was .64. The point-biserial correlation obtained for the correct response was .47. The index of discrimination obtained for option D ($r_{pbis} = -.07$) indicated that this option did not discriminate between high-scoring and low-scoring students. Option B discriminated sufficiently ($r_{pbis} = -.16$). However, this option, as well as option D, were chosen by a low proportion of students (.05, .07).

Item Revision. Revisions were guided by the functioning of option C ($r_{pbis} = -.41$), the option chosen by 25% of the students. In this option the change in contour was imbedded on the second half of the second beat in measure 3 (see Figure 66). For Test Version 7, the rhythm employed in measure 2 (Test Version 6) was replaced by a series of eighth notes, the contour of which differs from that of the correct response in

Table 21

Field Study 6: Indices of Difficulty and Discrimination

				Options				
Item	Task	Format	Index	A	B	C	D	Revise
1 (1)	follow notes	WR	p	.73*	.27			
			r_{pbis}	.45	-.45			
2 (2)	follow notes	WR	p	.50*	.50			
			r_{pbis}	.39	-.39			
3 (3)	follow notes	WR	p	.64*	.36			
			r_{pbis}	.58	-.58			
4 (4)	follow notes	WR	p	.36*	.64			
			r_{pbis}	.36	-.36			
5 (7)	home tone	WR	p	.43*	.57			
			r_{pbis}	.55	-.55			
6 (8)	home tone	WR	p	.62*	.38			
			r_{pbis}	.52	-.52			
7 (9)	read notes	MC	p	.47*	.15	.22	.16	
			r_{pbis}	.54	-.20	-.27	-.23	
8 (10)	read notes	MC	p	.09	.21	.54*	.15	
			r_{pbis}	-.20	-.28	.48	-.19	
9 (11)	contour	MC	p	.04	.88*	.08		
			r_{pbis}	-.26	.39	-.28		
10 (12)	contour	MC	p	.26	.45*	.29		
			r_{pbis}	-.24	.47	-.29		
11 (13)	read notes	MC	p	.64*	.05	.25	.07	x
			r_{pbis}	.47	-.16	-.41	-.07	
12 (14)	read notes	MC	p	.12	.14	.14	.60*	
			r_{pbis}	-.28	-.27	-.26	.56	

Note. * denotes correct response. MC = multiple-choice; WR = written response. Numbers in parentheses refer to item order in Field Study 5. For items 21 and 22, option D denotes responses which could not be categorized.

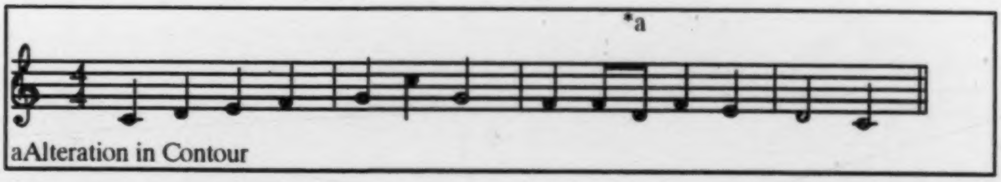
Table 21

Field Study 6: Indices of Difficulty and Discrimination

				Options				
Item	Task	Format	Index	A	B	C	D	Revise
13 (15)	sharps	MC	p	.34	.24	.42*		x
			r_{pbis}	.05	-.28	.19		
14 (16)	flats	MC	p	.47*	.28	.25		
			r_{pbis}	.33	-.19	-.19		
15 (17)	leap	MC	p	.34*	.66			
			r_{pbis}	.54	-.54			
16 (18)	step	WR	p	.16*	.84			x
			r_{pbis}	.39	-.39			
17 (19)	names	WR	p	.61*	.39			
			r_{pbis}	.57	-.57			
18 (20)	names	WR	p	.68*	.32			
			r_{pbis}	.59	-.59			
19 (21)	solfa	WR	p	.39*	.61			
			r_{pbis}	.68	-.68			
20 (22)	solfa	WR	p	.50*	.50			
			r_{pbis}	.62	-.62			
21 (23)	sequence	MC	p	.26	.19	.43*	.13	x
			r_{pbis}	-.17	-.18	.50	-.13	
22 (24)	sequence	MC	p	.41*	.28	.20	.11	x
			r_{pbis}	.38	-.17	-.05	-.29	

Note. * denotes correct response. MC = multiple-choice; WR = written response. Numbers in parentheses refer to item order in Field Study 5. For items 21 and 22, option D denotes responses which could not be categorized.

options B and D (see Figure 67). This revision was undertaken in an attempt to attract students away from option C to the other two incorrect options (B and D).



[a]Alteration in Contour

Figure 66. Item 11: Option C

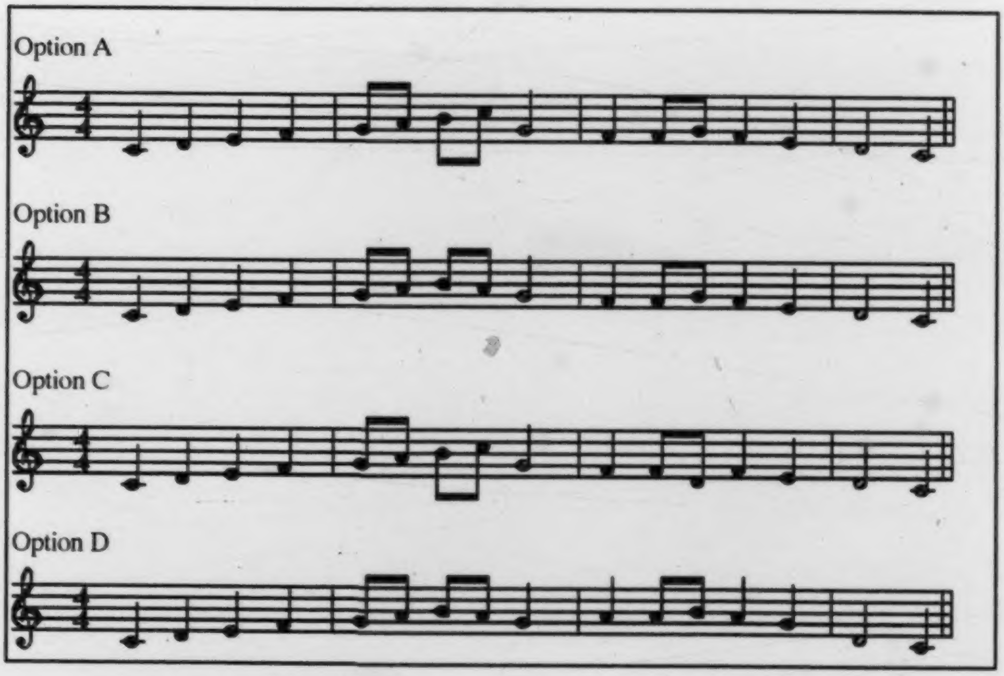


Figure 67. Item 11: Test Version 7.

Item 13

Item 13 was designed to assess a student's ability to identify the flats in an aural presentation of a melody.

Indices of Difficulty and Discrimination. The difficulty index for item 13 was .42. The point-biserial index of correlation for the correct response (option C) was .19. Point-biserial indices of discrimination obtained for options A and B were .05 and -.28, respectively. Thus, it appears that, in general, high-scoring students chose between options A and C in response to this item. More specifically, the high-scoring students dismissed the option in which the notes were unaltered and chose their response from between the two options in which the notes were altered by either sharps or flats. Considering that the point-biserial index of discrimination for option C (the correct response) was higher than that of option A (an incorrect alternative), it would appear that response was not entirely random.

Item Revision. Possible problems were noted with respect to the discrimination of option A. However, the item was not revised further pending information obtained in a subsequent test administration.

Item 16

Item 16 was designed to assess a student's ability to identify all the notes which moved by a step.

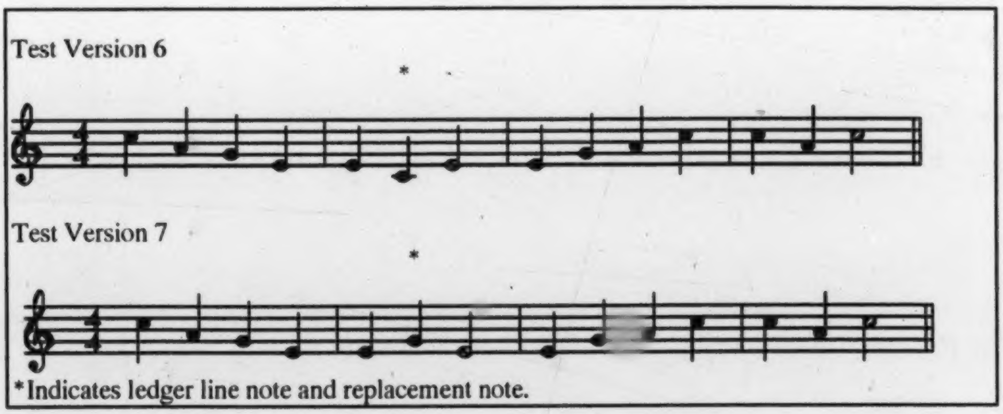
Indices of Difficulty and Discrimination. The difficulty index for item 16, .16, was below the suggested guideline of .20 . The point-biserial index of discrimination was .39.

Item Revision. It is possible that correct response was influenced by the structure of the melody employed for this item. More specifically, the melody employed in item 16 included a ledger line note below the staff. In revised form, the ledger line note was replaced by a note on the second line (see Figure 68).

Items 21 and 22

Items 21 and 22 were designed to assess a student's ability to identify a repeating melodic sequence.

Indices of Difficulty and Discrimination. The difficulty indices for items 21 and 22 were .43 and .41. For item 21, the point-biserial index of correlation for the correct response was .50 and all incorrect options discriminated negatively. For item 22, the



Test Version 6

Test Version 7

*Indicates ledger line note and replacement note.

Figure 68. Item 16: Melody employed in Test Versions 6 and 7.

point-biserial index of correlation for the correct response was .38. All incorrect responses discriminated negatively.

Item Revision. The difficulty students had in understanding these items is demonstrated by the number of responses that could not be categorized (option D). Although this option discriminated negatively, it must be noted that for items 21 and 22 respectively, 13% and 11% of the students did not know how to proceed with these questions. For the most part, rather than choosing the option in which the initial melodic sequence was repeated, these students circled portions of melodies among options. Based on this observation, a practise question was provided in an attempt to ensure that students understand how to proceed with these items.

Summary Test Statistics

The test scores were aggregated for each individual to produce a total score out of 22. For the sample of 214 students, the mean was 11.25 (51.1%), the median was 10.0 and the mode was 9.0. The scores ranged from 1 to 21, with a standard deviation of 5.0 (4.4%). The internal consistency, coefficient alpha (Cronbach, 1951), was .84; the corresponding SEM was 2.01. Taken together, the test as a whole was of medium difficulty. The scores were essentially symmetrically distributed around the center point.

The internal consistency exceeded the .80 value generally accepted for this type of measure (Anastasi, 1988; Ebel & Frisbie, 1991).

Test Validity

Students' Opportunity to Learn

For students involved in Field Study 6, opportunity to learn was examined by means of a questionnaire administered to each of the 7 music specialists.

Structure of the Questionnaire

The structure of the questionnaire was governed, in part, by the framework of the test. One item was randomly selected from each pair of items in Test Version 6. The selected items were 1, 3, 5, 7, 9, 12, 13, 15, 18, 19, and 22. Three questions were asked about each item. In addition, a fourth, more general question was asked about the teaching materials used by the teachers. The four questions were:

1. Characteristic of Melody: In your estimation, what is the extent to which students in your music program have had the opportunity to engage in musical activities which relate to the characteristic of melody described above?
   Scale:
   1 = full opportunity 2 = some opportunity 3 = no opportunity
2. Sample Test Item: In your estimation, what is the extent to which students in your program have had the opportunity to engage in musical activities similar to the requirements of the sample test item?
   Scale:
   1 = full opportunity 2 = some opportunity 3 = no opportunity
3. Time of Instruction: When did students last have the opportunity to engage in follow-up activities similar to the requirements of the sample test item?

Sample Descriptors (given by the teachers to best specify when students last engaged in related activities):

LC = last class LM = last month LW = last week LY = last year
NR = not recently CON = continuously

4. Instructional Materials: In brief, please describe the books and materials you use in lesson preparation.

Response:

A variety of written responses were obtained for this question.

For questions 1, 2, and 3 responses were entered in boxes provided on the questionnaire form. For question 4 responses were entered on the lines provided on the questionnaire form. Responses to questions 1, 2, 3, and 4 are summarized in Tables 22, 23, 24, and 25, respectively. An abbreviated form of the questionnaire, showing these 4 questions with reference to a single test item, is presented in Appendix N.

Summary of Responses

Characteristic of Melody (question 1). As summarized in Table 22, at least 6 teachers (> 85%) indicated that students had at least some opportunity to experience music in terms of the characteristics of melody identified in this study for items 1, 3, 5, 7, 9, 12, 15, and 18. In terms of item difficulty, with the exception of item 15, these were the easiest items. The difficulty of item 15, p = .34, suggests that while the students appeared to have experienced music in terms of the characteristic of melody assessed by this item, they were not able to identify notes that move by leap as represented by this item.

Three teachers indicated that students had no opportunity to engage in musical activities related to the characteristic of melody assessed by items 13 and 22. In this respect, an inconsistency in response was noted on the part of a teacher who indicated that students had no opportunity to engage in musical activities related to item 22 but, for

Table 22

Questionnaire Responses: Characteristic of Melody (question 1)

Item	Difficulty	Opportunity to Learn		
		Full	Some	None
9	.88	4	3	0
1	.73	4	3	0
18	.68	5	2	0
3	.64	5	2	0
12	.60	4	3	0
7	.47	2	5	0
5	.43	1	5	1
13	.42	1	3	3
22	.41	1	3	3
19	.39	1	4	2
15	.34	2	5	0
Total		30 (40%)	38 (49%)	9 (11%)

Table 23

Questionnaire Responses: Test Items (question 2)

Item	Difficulty	Opportunity to Learn		
		Full	Some	None
9	.88	5	2	0
1	.73	1	5	1
18	.68	5	2	0
3	.64	3	3	1
12	.60	0	6	1
7	.47	1	3	3
5	.43	0	6	1
13	.42	1	1	5
22	.41	1	3	3
19	.39	1	3	3
15	.34	1	5	1
Total		19 (25%)	39 (50%)	19 (25%)

Table 24

Questionnaire Responses: Time Period (question 3)

	School						
Item	1	2	3	4	5	6	7
9	LM	2M	LM	CON	LM	NG	CON
1	LW		LC	9M	LM	LW	3M
18	LC	3W	LC	1M	LM	LW	3M
3	LY		LC	9M	NR	LY	CON
12	LC		LM		LM		NG
7	LM	3W	2M	4M			
5	LM	2M	LM	CON	LM	NG	CON
13			LM	LY	4M		
22			LC	LY	3M		NG
19		7M		7M	LY	LY	NG
15	LW	4W	2M	CON	LM		NG

Note. LC = last class; LM = last month; LW = last week; LY = last year

M = month; W = week; NG = not given; NR = not recently; CON = continuously.

Table 25

Instructional Materials

	School						
Materials	1	2	3	4	5	6	7
Curriculum			x			x	
Musicanada	x				x		x
Music and You	x	x		x	x		x
Kodaly-Based Materials					x	x	
Orff-Schulwerk Materials		x	x	x	x		x
Folk Song Books	x					x	x
Recorder Books	x	x		x			

question 2, indicated that students had some opportunity to engage in musical activities similar to the requirements of the corresponding test item. Nonetheless, it would appear that students had limited experience with flats (item 13) and melodic sequences (item 22)—two of the more difficult items on the test.

Test Items (question 2). As summarized in Table 23, at least 6 teachers (> 85%) indicated that students had at least some opportunity to engage in musical activities similar to the requirements of test items 1, 3, 5, 9, 15, and 18. Five teachers indicated that students had at least some opportunity to engage in musical activities similar to the requirements of item 12. For the most part, these items were among the easiest on the test. However, as with question 1, performance on item 15 did not appear to reflect the students opportunity to engage in musical activities similar to the requirements of the item. Conversely, item 7, for which 3 teachers indicated that students had no opportunity to engage in musical activities similar to the requirements of the item, had a difficulty index of .47. Considering that students in all schools had at least some opportunity to experience music in terms of the characteristic of melody referenced to this item (see question 1), it may be that students responded to this item by applying musical knowledge acquired in other situations.

At least 3 teachers indicated that students had no opportunity to engage in musical activities similar to the requirements of items 13 ($p = .42$), 22 ($p = .41$), and 19 ($p = .39$). With the exception of item 15, these items were the most difficult items assessed on the questionnaire. In this respect, it was noted that item 5 ($p = .43$), an item in which students in at least 6 schools had some opportunity to engage in related musical activities, had a p-level similar to that of those items for which the students' opportunity to learn was more limited. It is possible that the difficulty of this item was related to the structure of the melody employed for the task. More specifically, in item 5 students were required to locate the home tone for a minor melody which concluded with the A- G# -A configuration. It is possible that the inclusion of an accidental (#) confused those students unfamiliar with this experience music in terms of the characteristics of melody identified in the analysis of the educational literature. In addition, responses indicated that, in general, difficulty levels were higher in those items for which at least 6 teachers specified that students had at least some opportunity to engage in musical activities similar to the requirements of the corresponding test items as compared to those items in which students' opportunity to learn was more limited. Furthermore, questionnaire responses showed that students engaged in follow-up activities similar to the requirements of the test items during the current school year, indicating that the test items developed in this study were referenced to instructional objectives in which grade 5 students gain experience.

Influence of Formal Music Instruction and Gender

The influence of private music instruction and gender upon performance on Test Version 6 was examined by means of analysis of variance. The students were requested to indicate whether they had received formal music instruction outside that received in school music classes (see Appendix O). Those students who indicated they had at least one year of formal music education fitting this description were designated as experienced, regardless of whether they were currently taking these lessons. Based on this criteria, 100 students were assigned to this group. Ten of the 100 students indicated that they had received private instruction on more than one instrument. Eighty-seven students played the piano, 13 students played an orchestral string instrument (violin or cello), and 3 students played the guitar. The bagpipes, clarinet, flute, and saxophone were each listed by one student as the instrument on which they had received formal instruction. Five students indicated that they had taken private voice lessons. Those students who had not participated in formal music classes beyond those offered at school were identified as musically inexperienced (N = 114). The distribution of students by experience and gender is reported in Table 26.

motive. In consideration of this, it was noted that in item 6 (p = .62) students were requested to locate the home tone for a major melody which concluded with a step-wise descent to the tonic.[1]

Time of Instruction (question 3). As summarized in Table 24, a variety of responses were submitted in relation to the time at which students had last engaged in follow-up activities similar to the requirements of the representative test items. The descriptor "last year" (LY), indicating that students had engaged in follow-up activities in the previous school year, was used 5 times. The remaining descriptors revealed that students had taken part in follow-up activities during the current school year. This demonstrates that the items developed in this study were referenced to activities in which grade 5 students gained experience.

Instructional Materials (question 4). As summarized in Table 25, music teachers involved in Field Study 6 used a variety of educational materials. Teachers in schools 3 and 6 indicated that they used curriculum guides. Teachers in schools 1, 2, 4, 5 and 7 used materials from the music series Musicanada and/or Music and You. Teachers in schools 5 and 6 used Kodaly-based materials and teachers in schools 2, 3, 4, 5, and 7 used materials based on the philosophy and principles of Orff-Schulwerk. Teachers in schools 1, 6, and 7 supplemented educational materials with content from folk song sources and teachers in schools 1, 2, and 4 used materials designed for teaching recorder.

Summary

For students in Field Study 6, opportunity to learn was examined by means of questionnaire responses obtained from each of the music specialists who participated in this study. Responses revealed that, to a large extent, students had at least some opportunity to

---

[1] With the exception of those items in which students were required to identify sharps or flats, the only accidental used in melodies composed for this study was the raised seventh degree of the harmonic minor scale.

Table 26

Cell Distribution

Gender	Experience		Total
	Experienced	Inexperienced	
Males	44	67	111
Females	56	47	103
Total	100	114	

A 2 x 2 (experience-by-gender) fixed effects analysis of variance was performed. Given that the design was non-orthogonal due to unequal sample sizes in each cell, the order of the factors (experience followed by gender) was determined *a priori* using a hierarchical model (Norusis, 1993). Results disclosed that there was a significant ($p < .01$) difference between the scores obtained by students in the experienced group as compared to students in the inexperienced group (see Table 27). There was no difference between males and females nor was there a significant interaction between experience and gender.

Table 27

Analysis of Variance: Total Score (Dependent Variable)

source	df	MS	F
-Experience (E)	1	1894.724	113.060 **
-Gender (G)	1	5.032	.300
Interaction (E x G)	1	.321	.019
Residual	210	16.759	
Total	213	25.443	

Note. ** p< .01.

These findings confirmed the expected results. First, as demonstrated in the review of the literature, several prominent researchers and writers in the field of music education refer to music achievement in terms of accomplishments acquired as a result of formal educational experiences (Boyle, 1992; Boyle & Radocy, 1987; Colwell, 1970; Gordon, 1965; Radocy, 1989). In consideration of this definition, one would expect that students with a more extensive experiential background in music would outperform students whose only formal instruction was acquired in school music classes. Second, it was expected that there would be no difference between the test performance of males and females. Of the 10 tests in which this aspect of performance was examined in the literature review (see Chapter 2), the difference in performance between males and females was found to be significant for only two measures—The Gaston Test of Musicality (Gaston, 1957) and the Farnum Music Test (1969). For the most part, where differences were found, they were not sufficient to warrant the formulation of separate norms.

Factor Structure

Overview

Factor analytic techniques were used to summarize the empirical relationships among the test items developed in this study. As summarized in Gorsuch (1983), the aim of factor analysis is to "summarize the interrelationships among the variables [in this study, test items] in a concise but accurate manner as an aid in conceptualization" (p. 2). This is done by reducing the number of variables to a smaller number of factors, each of which represents a distinct area of generalization (Gorsuch, 1983).

For purposes of this study, the number of factors was determined using the Kaiser-Guttman rule (Kaiser, 1960), the scree test (Cattell, 1966), the presence of simple structure (Gorsuch, 1983), and the interpretation of results (Messick, 1989b). As Messick (1989b) explained, "so-called 'process' interpretations are inferences drawn from test outcomes about possible processes that might plausibly have produced the correlated results and, as such, they constitute provisional constructs" (p. 52). In this respect, as presented in the current study, the interpretive review of factor loadings encompasses several hypotheses, all of which are subject to further empirical testing and re-interpretation (Messick, 1989b).

Factor interpretations were based upon the identification of factor loadings with sufficient magnitude "to assume that a relationship exists between the variable and the factor" (Gorsuch, 1983, p. 208). According to Gorsuch (1983), only rough guidelines can be given to determine salient loadings. Among the methods of determining the lower bound is the criteria of meaningfulness. In Gorsuch's view "this may be the reason for the popularity of an absolute value of .3 as the minimum loading for interpretation" (p. 210). This is the value used in the current analysis.

Method

Using a principal components analysis, 7 factors were revealed using the Kaiser-Guttman rule of eigenvalues over 1 (Kaiser, 1960). Two factors were revealed in the scree test (Cattell, 1966). Orthogonal solutions (varimax) were generated, extracting from 2 to 7 factors. None of these solutions produced a pattern matrix which approached simple structure. Therefore, oblique solutions (direct obliman) were generated, again extracting from 2 to 7 factors. The three-factor solution produced a pattern matrix which approached simple structure. An abbreviated form of the pattern matrix, showing factors loadings with absolute values exceeding .3, is presented in Table 28; the complete pattern matrix is provided in Appendix P. The correlations among the 3 factors are reported in Table 29.

Of the 22 items, 17 items (1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 15, 16, 17, 18, 19, 21, and 22) loaded on a single factor. Items 11, 13, 14, and 20 were complex variables. Of these, items 13 and 20 loaded on factors 1 and 2; item 11 loaded on factors 1 and 3; and item 14 loaded on factors 2 and 3. No salient loadings were produced for item 8.

Description of the Factors

Three factors were identified in this analysis: factor 1: aural/visual; factor 2: visual; and factor 3: contour. Each factor is described in terms of those items with salient loadings on a single factor.

Table 28

Factor Analysis: Rotated to Obliman Solution (Pattern Matrix) to Approximate Simple Structure

	Factors		
Items	1 Aural/Visual	2 Visual	3 Contour
1	.50917	-	-
2	.47537	-	-
3	.51044		
4	.63720	-	-
5	.59689	-	-
6	.46990	-	-
7	.48744	-	-
8	-	-	-
9	-	-	.67209
10	-	-	-.45852
11	.37624	-	-.36162
12	-.41125	-	-
13	.42460	.31069	-
14	-	-.42159	-.30998
15	-	-.58801	-
16	-	-.62876	-
17	-	-.68577	-
18	-	-.67065	-
19	-	-.71279	-
20	.32064	-.52069	-
21	-	-	-.41562
22	-	-	-.51436

Note. - = < .3.

Table 29

Factor Correlation Matrix

	Factor 1	Factor 2	Factor 3
Factor 1	1.00000		
Factor 2	-.22069	1.00000	
Factor 3	-.06223	.05329	1.00000

Factor 1: Aural/Visual. Items 1, 2, 3, 4, 5, 6, 7, and 12, loaded on factor 1. All of these items were presented in both visual and aural form—hence the label assigned to this factor. Items associated with five tasks were included in this group: both items in which students identified the last note sung (items 1 and 2); both items in which students identified the silent notes (items 3 and 4); both items in which students identified the home tone of a given melody (5 and 6); the first of two items in which students identified the one measure of a given melody in which the aural presentation differed from the notational representation (item 7); and the second of two items in which students identified the notational equivalent of an aural melody when presented with four options in notational form (item 12). Considering that the first 14 items were presented both aurally and visually, it was reasonable to observe that these items clustered in terms of this method of presentation.

Factor 2: Visual. Five of the 8 items presented in visual form only (items 15, 16, 17, 18, and 19) loaded on factor 2. Items representing three tasks were included in this group: both items in which students identified notes which moved by leap or step (items 15 and 16); both items in which students identified notes by their absolute letter names (items 17 and 18); and the first of two items in which students identified notes by their solfa syllables (item 19). The emergence of this factor seemed to confirm the findings of Field Studies 4 and 5.

Factor 3: Contour. As demonstrated in the factor analytic data, the importance of melodic contour was revealed in items 9, 10, 21, and 22. Items representing two tasks were included in this group: both items in which students identified the overall contour of an aurally-presented melody when presented with three notational representations (items 9 and 10); and both items in which students identified a repeating melodic sequence (items 21 and 22). Within this group, each pair of items differed in terms of method of presentation: items 9 and 10 were presented in both aural and visual form while items 21 and 22 were presented in visual form only. Thus, for these items, it seems that contour information emerged as a central component of musical understanding. This corresponds with the research of Dowling (1991) which demonstrated that melodies are heard as integrated wholes in which recognition is influenced by both contour, interval, and tonality.

Complex Variables

Items 13 and 20 loaded on factors 1 and 2; item 11 loaded on factors 1 and 3; and item 14 loaded on factors 2 and 3. Of these, items 11, 13, and 14 were presented in aural and visual mode and item 20 was presented in visual mode only.

Item 11. In item 11, students were requested to identify the notational equivalent of an aural melody when presented with four options in notational form. This item, which was presented both aurally and visually, loaded on both the aural/visual factor (factor 1) and the contour factor (factor 3). It would appear that the importance of the contour factor is due to the structure of the melody (see Figure 69: frame 1). As illustrated in Figure 69, the melody for the correct response opens with an ascending motive beginning on middle 'C' and concludes with a descending motive ending on middle 'C'. The overall rise and fall of this melody is similar to that employed in item 9, an item which loaded solely on the contour factor (see Figure 69: frame 2). This is in contrast with item 12, an item of similar structure to 11 which loaded on factor 1 only. As illustrated in Figure 69 (frame 3), the melody in item 12 does not employ an ascending or descending contour such as that found in items 9 and 11.

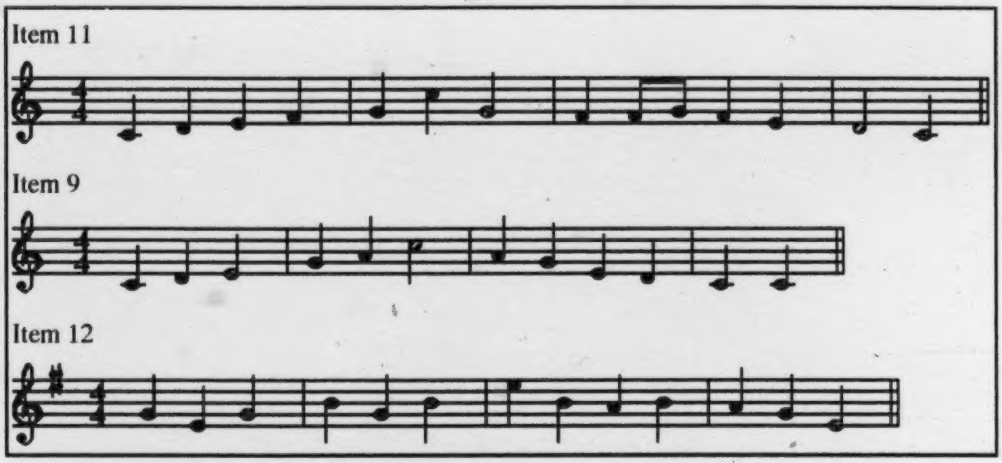


Figure 69. Items 11, 9, and 12: Correct response.

Item 13. In item 13 students were requested to identify the sharps in a given melody. This item loaded on both factor 1 (aural/visual) and factor 2 (visual). Considering that students were requested to listen to a given melody, it was expected that this item would load on factor 1 (aural/visual). The emphasis on the visual factor may be due, in part, to the structure of the item. More specifically, with the exception of the sharps and flats, each melody employed the same notation. Thus, in order to identify the notation which was the same as the melody they heard, students were required to observe that all three melodies employed the same notes (emphasis on the visual) as well as to discern whether the melody they heard employed either flats or sharps (emphasis on the aural/visual).

Item 14. In item 14 students were requested to identify the flats in a given melody. This item loaded on factor 2 (visual) and factor 3 (contour). As with item 13, this item may have loaded on factor 2 because, with the exception of the sharps and flats, all three options employed the same notation. It is possible that the emphasis on factor 3 is due to the structure of the melody employed in this task (see Figure 70). As illustrated in this figure, both measures 1 and 3 employed the same notation—hence the emphasis on contour revealed in the pattern matrix.

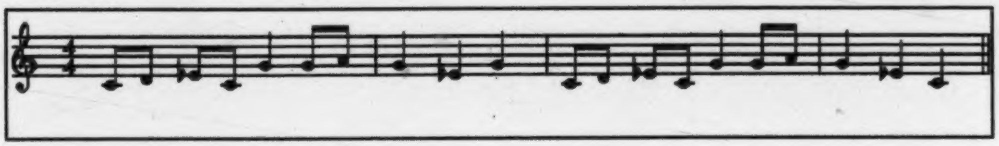


Figure 70. Item 14: Correct response.

Item 20. In item 20, students were requested to identify notes by their corresponding solfa syllables. This item loaded on factor 1 (aural/visual) and factor 2 (visual). Considering that this item was presented in visual form only, one would expect it to load on factor 2. The emphasis on factor 1 implies that, in response to this item, students sang this melody in their heads. This explanation cannot be confirmed from the available data and is therefore subject to further investigation.

Nonloading Item

Item 8 did not load on any of the three factors identified in the analysis. A rationale for this observation was formulated by comparing the requirements of item 8 with those of item 7, an item of similar structure which loaded exclusively on factor 1. Both items 7 and 8 were designed to assess a student's ability to identify the one measure of a given melody in which the aural presentation differed from the notational representation. In item 7, the notation differed from the aural representation in terms of both contour and interval in the first measure.[2] For item 8, the notation differed from the aural representation in terms of interval on two notes in the third measure. In this respect, the difference between the aural and visual representations of the melody were more subtle than those of the previous item. This may account, in part, for a difference in the way in which students responded to these items. Beyond this, the underlying functioning of item 8 cannot be inferred from the factor analytic results.

---

[2] For item 7, the variation in contour between the aural and notational representation is limited to one beat (two eighth notes). As a factor, contour loads on those items for which a melody is characterized by a specific melodic contour, the duration of which extends for at least an entire measure.

Factor Correlation

Examination of the correlations among factors revealed a weak, but negative, correlation between factors 1 and 2. Factor 3 was essentially uncorrelated with both factors 1 and 2.

In correspondence to the method of administration, items loading exclusively on factor 1 were presented in a combined aural/visual mode while items loading exclusively on factor 2 were presented in visual mode only. Considering the low negative correlation between these factors, it would appear that the way in which students responded to the items presented in the aural/visual mode differed from the way in which students responded to the items presented in visual mode only.

Musically, the items loading on factor 3 were characterized by melodies with specific contours, the durations of which extended across four-beat sequences (items 21 and 22) or entire phrases (items 9 and 10). Considering that the melodic structure of items that loaded on factor 3 differed from the melodic structure of the remaining items, this factor was essentially uncorrelated with either factor 1 or factor 2.

Summary

Factor analytic techniques were used to summarize the empirical relationships among the test items developed in this study. An oblique solution approaching simple structure produced 3 factors: factor 1: aural/visual; factor 2: visual; and factor 3: contour.

Considering that the first 14 items were presented both aurally and visually, it was reasonable to observe 8 of these items cluster solely in terms of factor 1 (aural/visual).

Five of the 8 items presented in visual form only (items 15 , 16, 17, 18, and 19) loaded solely on factor 2 (visual). The emergence of this factor seemed to confirm the findings of Field Studies 4 and 5 (Chapter 7: Mode of Presentation). Items 9, 10, 21, and 22 loaded solely on Factor 3 (contour). This factor identified melodies with specific contours, the duration of which extended for at least an entire measure. Four complex variables were identified in this analysis. Of these, items 13 and 20 loaded on factors 1 and 2; item 11 loaded on factors 1 and 3; and item 14 loaded on factors 2 and 3. The low correlations among factors reveals that, to a large extent, students responded differently to each group of items.

Chapter Summary

Test Version 6, comprised of a total of 22 items, was administered to 214 grade 5 students. All students attended schools within School District 1, 2, or 3 and all students were taught by a music specialist. Based on the indices of difficulty and discrimination obtained in this study, 17 items were not revised (items 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 17, 18, 19, and 20). Item 11 was revised in terms of rhythm and item 16 was revised in terms of melody. For items 21 and 22, a practise question was provided in an attempt to ensure that students understand how to proceed with these items. Although data obtained for item 13 was not fully satisfactory, this item remained in its present form pending further information obtained in subsequent test administration.

Students' opportunity to learn was examined by means of questionnaire responses obtained from each of the music specialists involved in this study. Responses indicated that, in general, difficulty levels were higher in those items for which at least 6 teachers specified that students had at least some opportunity to engage in musical activities similar to the requirements of the corresponding test items as compared to those items in which the students' opportunity to learn was more limited. Furthermore, questionnaire responses showed that students were engaged in follow-up activities similar to the requirements of the test items during the current school year, indicating that the test items developed in this study were referenced to activities in which grade 5 students gain experience.

For purposes of this study, students who indicated they had received at least one year of formal musical training outside of school were designated as experienced, regardless of whether they were currently taking private lessons. Using this criteria, 100 students were assigned to this group. Analysis of variance disclosed that there was a significant difference between the scores obtained by students in the experienced group as compared to students in the inexperienced group. There was no difference between males and females nor was there a significant interaction between experience and gender.

Factor analytic techniques were used to summarize the empirical relationships among test items developed in this study. An oblique solution approaching simple structure produced 3 factors: factor 1: aural/visual; factor 2: visual; factor 3: contour. Weak correlations among factors indicated that, to a large extent, students responded differently to each group of items identified as a factor.

CHAPTER 9

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The purposes of this study were (1) to identify the various ways in which music literacy (the ability to read and/or write tonal melodies in traditional Western music notation) is addressed in teaching materials currently employed in school music programs at the grade 5 level and, using this information, (2) to devise pencil-and-paper test items by which students might demonstrate their performance in terms of these items. The first question was addressed by analyzing selected educational resources in which music literacy was viewed as a fundamental goal of instruction. The second question was addressed in three stages: 1. the development of the initial test items based on the professional judgement of the test developer; 2. the assessment of the relevance and representativeness of these items with reference to the specified content domain based on the professional judgement of subject-area specialists; and, 3. the examination of student response based on data obtained in six field studies, each of which employed a revised version of the test. Two underlying issues were addressed in test administration: first, the function of the test items as revealed in classical item analysis (Field Studies 1, 3, 5, and 6) and second, the strategies employed by students when solving these items as demonstrated in the data obtained through the analysis of student protocols (Field Studies 2 and 4).

Summary of Findings

The process of item development has not been examined in former investigations which study the development of test items in the subject area of music. Generally, in preliminary studies, test developers administer a greater number of items than that desired for the final form of the test. Based on these data, those items that do not satisfy a predetermined criteria for item difficulty and discrimination are deleted from the study. From the perspective of test construction, this procedure must be used with caution because the deletion of items changes the original specifications of the content domain (Cronbach, 1971). Findings from this study demonstrate that it is possible to improve items through the process of item revision.

While there are many specific findings revealed in Chapters 3 to 8, they can be collated and summarized in terms of the following five general findings:

1. Data revealed that the overall functioning of items was improved through the process of item revision. Revisions were assigned to five categories: directions to students, melody, rhythm, target note, and practise question.
2. Data revealed that student response was influenced by the musical structure of the item. Characteristics of melodic structure identified in this study included contour, interval (steps, leaps, repeats), tonality, and ledger line.
3. Data obtained in Field Study 4 revealed that items referenced to four tasks (the identification of steps or leaps, the identification of solfa syllables, the identification of absolute note names, and the recognition of sequential patterns) could be answered correctly by attending to the notation in the absence of the aural stimulus.
4. Based on the data obtained in Field Study 4 (Mode of Presentation), Test Versions 5 and 6 were arranged in two main sections. The first section contained the items which were presented both aurally and visually and the second section included the items which were presented visually. Considering the analysis of the factor structure (see Chapter 8), it would seem that the sequential ordering of items used in the final two field studies is appropriate for the items developed in this study.
5. Factor analytic data revealed that items clustered in three groups: factor 1: aural/visual; factor 2: visual; and, factor 3: contour. Seventeen of the 22 items loaded on a single factor. Items 13 and 20 loaded on factors 1 and 2; item 11 loaded on factors 1 and 3; and item 14 loaded on factors 2 and 3. No salient loadings were produced for item 8.

Limitations

1. The characteristics of melody were each assessed by two test items. As such, the items do not represent a comprehensive measure of each characteristic of melody identified in the analysis of the educational literature.
2. This study examines music proficiencies in terms of the ability to discriminate written notation. As such, the test items developed in this study are not an all-encompassing measure of music ability.
3. Findings are based on the data obtained from the students involved in the current study. Further research is needed to examine the applicability of these results beyond this parameter.

Implications for Future Research

Two data sources were consulted with reference to the administration of test items to students: 1. indices of difficulty and discrimination obtained from group administration; and 2. student protocols obtained from individual interviews. With regard to the data obtained through group administration, interpretation extended beyond reporting indices of difficulty and discrimination to explaining why specific indices may have been obtained. Item revisions were documented, thus justifying how tasks were altered in consideration of available data and suggesting how these alterations may have influenced the magnitude of indices obtained in subsequent administrations. Protocol data provided evidence of the strategies which students employed as they solved the problems presented in this study. Use of this research methodology in the present study demonstrated its applicability to the subject area of music and will hopefully influence the structure of future studies.

Due, in part, to the small number of items developed in this study, all preliminary findings are subject to further investigation. The construction of additional items may lead to the development of item pools. Convergent and divergent evidence of construct validity may be revealed by correlating the results obtained through the administration of these items with the outcomes obtained from other forms of data collection. In this respect, researchers must recognize that test construction relies at most on crude theory-sketches rather than consolidated theories (Cronbach, 1989). Nonetheless, findings of the current investigation point the way for future research, the results of which may eventually provide a theoretical foundation for construct interpretation. Listed below are three more specific suggestions for such research:

1. Items which assessed a student's ability to identify major and minor seconds and items which assessed a student's ability to notate aurally presented phrases were deleted from the test. Further research is needed to determine whether these items belong to the content domain as originally delineated.
2. It was found that items referenced to four tasks could be answered correctly by attending to the notation in the absence of the aural stimulus. Further research is needed to establish whether this finding was influenced by the structure of the items and/or the musical background of the students involved in the study.
3. Further research is needed to examine the extent to which indices of difficulty and discrimination obtained in the current studies were influenced by the structure of the melodies employed for specific tasks.

Implications For Practise

Many specific implications for practise were revealed in Chapters 3 to 8. The principal issues are summarized below:

1. Revisions to directions to students emphasized the importance of constructing directions which explain the specifications of the items without providing clues for test-wise students. Teachers are advised that, when constructing their own test items, directions to students must be worded carefully in order to describe clearly the specifications of the tasks, but without supplying students with clues to the answers.

2. Music teachers define repeated notes as notes repeated in succession on the same line or space of the staff. This was not the meaning assigned by students (see Field Study 1). Music teachers are advised to describe and clarify the meaning they assign to this term in order to avoid possible confusion on the part of their students.
3. Data obtained in Field Studies 2 and 4 revealed the feasibility of asking students to think aloud while answering selected items. Teachers may use similar questioning techniques to gain insights into how their students approach musical tasks.
4. The difficulty of items was influenced by the presence of ledger-line notes below the staff. In consideration of this, music teachers are advised that students may have more difficulty reading music which contains ledger-line notes than music in which all the notes are positioned within the musical staff. This observation may influence the melodic structure of materials used for sight-reading activities.
5. The ability to assign letter names or solfa syllables to given notes depends on the structure of the melody employed for the task. In consideration of this, teachers are advised that students may have more difficulty supplying letter names or solfa syllables to notes which move in disjunct motion as compared to notes which move in conjunct motion. Again, this observation may influence the melodic structure of materials used for sight-reading activities.

References

Abeles, H. F., Hoffer, C. R., & Klotman, H. R. (1994). *Foundations of music education* (2nd ed.). New York: Schirmer.

Afflerbach, P. & Johnson, P. (1984). Research methodology on the use of verbal reports in reading research. *Journal of Reading Behavior, 16*(4), 307-321.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Andrews, F. M., & Deihl, N. (1970). Development of a technique for identifying elementary school children's musical concepts. *Journal of Research in Music Education, 18*(3), 214-222.

Apfelstadt, H. (1984). Effects of melodic perception instruction on pitch discrimination and vocal accuracy of kindergarten children. *Journal of Research in Music Education, 32*(1), 15-24.

Bartlett, J. C., & Dowling, W. J. (1980). The recognition of transposed melodies: A key distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 501-515.

Bartlett, J. C., & Dowling, W. J. (1988). Scale structure and similarity of melodies. *Music Perception, 5*(3), 285-314.

Bean, K. L. (1938). An experimental approach to the reading of music. *Psychological Monographs, 50*(whole number 226).

Bean, K. L. (1959). Review of The Farnum Music Notation Test. In O. K. Buros (Ed.), *The fifth mental measurements yearbook* (pp. 380-381). Highland Park, NJ: Gryphon.

Bentley, A. (1966). *Musical ability in children and its measurement.* London: George G. Harrap.

Boltz, M. (1989a). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 749-761.

Boltz, M. (1989b). Rhythm and 'good endings': Effects of temporal structure on tonality judgments. *Perception and Psychophysics, 46,* 9-17.

Boltz, M. (1991). Some structural determinants of melody recall. *Memory and Cognition, 19*(3), 239-251.

Boltz, M. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception and Psychophysics,* 53(6), 585-599.

Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology, 18,* 389-431.

Boyle, J. D. (1992). Evaluation of music ability. In R. Colwell (Ed.), *Handbook of research on music teaching and learning: A project of the Music Educators National Conference* (pp. 247-265). New York: Schirmer.

Boyle, J. D., & Radocy, R. E. (1987). *Measurement and evaluation of musical experiences.* New York: Schirmer.

Brooks, P. L, Kovacs, B. A., & Trotter, M. M. (1983). *Musicanada (Grade 4): Teacher's Guide.* . Toronto: Holt, Rinehart & Winston.

Brooks, P. L, Kovacs, B. A., & Trotter, M. M. (1984a). *Musicanada (Grade 6): Teacher's Guide.* . Toronto: Holt, Rinehart & Winston.

Brooks, P. L, Kovacs, B. A., & Trotter, M. M. (1984b). *Musicanada (Grade 5): Teacher's Guide.* . Toronto: Holt, Rinehart & Winston.

Buckton, R. A. (1977). A comparison of the effects of vocal and instrumental instruction on the development of melodic and vocal abilities in young children. *Psychology of Music, 5*, 36-47.

Butler, D. (1983). The initial identification of tonal centres in music. In D. Rogers & J. A. Sloboda (Eds.), *The acquisition of symbolic skills* (pp. 251-261). New York: Plenum.

Butler, D. (1992). *The musician's guide to perception and cognition.* New York: Schirmer.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research,* 1(2), 245-276.

Choksy, L. (1988). *The Kodaly method: Comprehensive music education from infant to adult* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Colwell, R. (1968). *Music achievement tests 1 and 2.* Chicago: Follett Educational Corporation.

Colwell, R. (1970a). *Music achievement tests 3 and 4.* Chicago: Follett Educational Corporation.

Colwell, R. (1970b). *The evaluation of music teaching and learning.* Englewood Cliffs, NJ: Prentice-Hall.

Colwell, R. (1982). Evaluation in music: Perspicacious or peregrine. In R. Colwell (Ed.), *Symposium in music education: A festschrift for Charles Leonhard* (pp. 157-195). Urbana-Champaign, IL: University of Illinois.

Colwell, R. (1991). Evaluation. In R. Colwell (Ed.), *Basic concepts in music education II* (pp. 247-278). Niwot, CO: University Press of Colorado.

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music.* Chicago: University of Chicago.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* New York: Holt, Rinehart, & Winston.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16, 297-334.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443-507). Washington, DC: American Council on Education.

Cronbach, L. J. (1989). Construct validation after thirty years. In R. L. Linn (Ed.), *Intelligence* (pp. 147-171). Urbana, IL: University of Illinois.

Croonen, W. L. M. (1994). Effects of length, tonal structure, and contour in the recognition of tone series. *Perception & Psychophysics,* 55(6), 623-632.

Croonen, W. L. M., & Kop, P. F. M. (1989). Tonality, tonal scheme, and contour in delayed recognition of tone sequences. *Music Perception, 7*(1), 49-68.

Cross, I., Howell, P., & West, R. (1983). Preferences for scale structure in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 444-460.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 869-883.

Cuddy, L. L., Cohen, A. J., & Miller, J. (1979). Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology, 33,* 148-156.

Damschroder, D. (1992). *Foundations of music and musicianship.* New York: Schirmer.

Davidson, L., McKernon, P., & Gardner, H. (1981). The acquisition of song: A developmental approach. In *Documentary report of the Ann Arbor symposium: National symposium on the applications of psychology to the teaching and learning of music* (pp. 301-315). Reston, VA: Music Educators' National Conference.

Davidson, L. (1985). Preschool children's tonal knowledge: Antecedents of Scale. In J. Boswell (Ed.), *The young child and music: Contemporary principles in child development and music education: Proceedings of the music in early childhood conference* (pp. 25-39). Reston, VA: Music Educators' National Conference.

Deutsch, D. (1969). Music recognition. *Psychology Review, 76,* 300-307.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception and Psychophysics, 28*(5), 381-389.

Deutsch, D. (1982). *The psychology of music.* New York: Academic.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*(4), 341-354.

Dowling, W. J. (1982). Melodic information processing and its development. In D. Deutsch (Ed.), *The psychology of music* (pp. 413-429). New York: Academic.

Dowling, W. J. (1986). Context effects on melody recognition: Scale-step versus interval recognition. *Music Perception, 3*(3), 281-296.

Dowling, W. J. (1988). Tonal structure and children's early learning of music. In J. Sloboda (Ed.), *Generative processes in music* (pp. 113-128). Oxford: Clarendon.

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception and Psychophysics, 50*(4), 305-313.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. New York: Academic.

Drake, R. (1946). Aptitudes and aptitude testing. In P. L. Harriman (Ed.), *Encyclopedia of Psychology* (pp. 38-44). New York: Philosophical Library.

Drake, R. (1957). *Manual for the Drake Music Aptitude Test* (2nd ed.). Chicago: Science Research Associates.

Duell, O., & Anderson, R. C. (1967). Pitch discrimination among primary school children. *Journal of Educational Psychology, 58,* 315-318.

Ebel, R. L. (1956). Obtaining and reporting evidence on content validity. *Educational and Psychological Measurement,* 16, 269-281.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Edworthy, J. (1985). Interval and contour in melody processing. *Music Perception, 2*(3), 375-388.

Elementary music curriculum guide. (1989). Edmonton, AB: Alberta Education, Curriculum Support Branch.

Ericsson, K. A. (1987). Theoretical implications from protocol analysis on testing and measurement. In R. Ronning, J. C. Conoley, J. A. Glover, & J. C. Witt (Eds.), *Buros-Nebraska symposium on measurement and testing: Vol. 3. The influence of cognitive psychology on testing* (pp. 191-228). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87* (3), 215-251.

Ericsson, K. A., & Simon, H. A. (1985). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT.

Farnsworth, P. (1969). *The social psychology of music* (2nd ed.). Ames, IA: Iowa State University.

Farnum, S. E. (1969). *Farnum Music Test.* Riverside, RI: Bond.

Fiske, H. E. (1982). Chronometric analysis of selected pattern discrimination tasks in music listening. *Psychology of Music, 10*(1), 37-45.

Franklin, E. (1977). An experimental study of text notation. *Bulletin of the Council For Research in Music Education, 50,* 18-20.

Gaston, E. T. (1957). *A test of musicality* (4th ed.). Lawrence, KS: Odell's Instrumental Service.

George, W. E. (1980). Measurement and evaluation of musical behavior. In D. Hodges (Ed.), *Handbook of music psychology* (pp. 291-392). Lawrence, KS: National Association for Music Therapy.

Geringer, J. M. (1983). The relationship of pitch-matching and pitch discrimination abilities of pre-school and fourth grade students. *Journal of Research in Music Education, 31*(2), 93-99.

Goolsby, T. (1989). Computer applications to eye movement research in music reading. *Psychomusicology, 8,* 111-126.

Gordon, E. E. (1965). *Music aptitude profile.* Boston: Houghton Mifflin.

Gordon, E. E. (1971a). *Iowa Test of Music Literacy.* Iowa City, IO: Bureau of Educational Research and Service.

Gordon, E. E. (1971b). *The psychology of music teaching.* Englewood Cliffs, NJ: Prentice-Hall.

Gordon, E. E. (1979a). Developmental music aptitude as measured by the primary measures of music audiation. *Psychology of Music, 7*(1), 42-49.

Gordon, E. E. (1979b). *Primary measures of music audiation.* Chicago: G.I.A.

Gordon, E. E. (1982). *Intermediate measures of music audiation.* Chicago: G.I.A.

Gordon, E. E. (1984). *Learning sequences in music: Skill, content, and patterns.* Chicago: G.I.A.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Green, G. (1990). The effect of vocal modeling on pitch-matching accuracy of elementary schoolchildren. *Journal of Research in Music Education, 38*(3), 225-231.

Greenberg, M. (1976). Research in early childhood education: A survey with recommendations. *Bulletin of the Council For Research in Music Education, 45,* 1-20.

Gregory, T. B. (1972). The effects of rhythmic notation variables on sight-reading errors. *Journal of Research in Music Education, 20*(4), 462-468.

Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York: MacMillan.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communications and Technology Journal, 29*(2), 75-91.

Hackett, P. & Lindeman, C. (1988). *The musical classroom: Backgrounds, models, and skills for elementary teaching* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hair, H. I. (1977). Discrimination of tonal direction on verbal and nonverbal tasks by first-grade children. *Journal of Research in Music Education, 25*(3), 197-210.

Halpern, A. R., & Bower, G. H. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology, 95*(1), 31-50.

Hargreaves, D. J. (1986). *The developmental psychology of music.* Cambridge: Cambridge University.

Hedden, S. K. (1982). Prediction of music achievement in the elementary school. *Journal of Research in Music Education, 30*(1), 61-68.

Hermanson, L. W. (1972). An investigation of the effect of timbre on simultaneous vocal pitch acuity of young children (Doctoral Dissertation, Columbia University, 1971). *Dissertation Abstracts International, 32,* 3558A.

Hodges, D. A. (1992). The acquisition of music reading skills. In R. Colwell (Ed.), *Handbook of research on music teaching and learning: A project of the Music Educators National Conference* (pp. 466-471). New York: Schirmer.

Hoffer, M. L. & Hoffer, C. R. (1987). *Music in the elementary classroom: Musicianship and teaching.* San Diego, CA: Harcourt Bace Jovanovich.

Hudson, A. N. (1986). The development of an achievement test in basic musical concepts for primary grades (Doctoral dissertation, Indiana University, 1986). *Dissertation Abstracts International, 47,* 4315A.

Hufstader, R. A. (1977). An investigation of a learning sequence of music listening skills. *Journal of Research in Music Education, 25*(2), 184-196.

Jersild, A. T., & Bienstock, S. F. (1934). A study of the development of children's ability to sing. *Journal of Educational Psychology, 25,* 481-503.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics, 41,* 621-34.

Jones, M. R. (1993). Dynamics and musical patterns: How do melody and rhythm fit together? In T. J. Tighe & W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm* (pp. 67-92). Hillsdale, NJ: Lawrence Erlbaum.

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics, 32,* 211-218.

Joyner, D. R. (1969). The monotone problem. *Journal of Research in Music Education, 17*(1), 115-124.

Kail, R. V. & Bisanz, J. (1982). Cognitive strategies. In C. R. Puff (Ed.), *Handbook of research methods in human memory and cognition* (pp. 229-255). New York: Academic.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 1,* 141-151.

Kennedy, M. (1980). *The concise Oxford dictionary of music.* London: Oxford University.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11,* 346-374.

Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception, 1,* 28-62.

Krumhansl, C. L. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology, 42,* 277-303.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579-594.

Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition, 10,* 243-251.

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in schoolchildren.* Chicago: University of Chicago.

Kwalwasser, J. (1955). *Exploring the musical mind.* New York: Coleman-Ross.

Kwalwasser, J., & Dykema, P. W. (1930). *K-D Music Tests: Manual of Directions.* New York: Carl Fischer.

Larson, W. S. (1959). Review of The Farnum Music Notation Test. In O. K. Buros (Ed.), *The fifth mental measurements yearbook* (pp. 380-381). Highland Park, NJ: Gryphon.

LeCompte, M. D. & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research, 52* (1), 31-60.

Lehman, P. R. (1968). *Tests and measurement in music.* Englewood Cliffs, NJ: Prentice-Hall.

Lehman, P. R. (1972). Review of Music Achievement Tests. In O. K. Buros (Ed.), *The seventh mental measurements yearbook* (Vol. 1, pp. 526-527). Highland Park, NJ: Gryphon.

Lehman, P. R. (1978). Review of Iowa Tests of Music Literacy. In O. K. Buros (Ed.), *The eighth mental measurements yearbook* (Vol. 2, p. 183). Highland Park, NJ: Gryphon.

Lehman, P. R. (1993). Implications of national standards. *Music Educators Journal, 80* (3), 25-28.

Linn, R. L., & Gronlund, N. E. (1995). *Measurement and assessment in teaching* (7th ed.). Englewood Cliffs, NJ: Merrill.

Long, P. A. (1977). Relationships between pitch memory in short melodies and selected factors. *Journal of Research in Music Education, 25*(4), 272-82.

Lundin, R. W. (1985). *An objective psychology of music* (3rd ed.). Malabar, FL: R. E. Krieger.

Mainwaring, J. (1931). Experiments on the analysis of cognitive processes involved in musical ability and in musical education. *British Journal of Educational Psychology, 1,* 180-203.

Mathews, W. S. B., & Liebling, E. (1973). *Dictionary of music.* Cincinatti: John Church. (Original work published 1896).

McCarthy, J. F. (1980). Individualized instruction, student achievement, and dropout in an urban elementary instrumental music program. *Journal of Research in Music Education, 23*(1), 59-69.

McDonald D. T., & Ramsey, J. H. (1979). A study of musical auditory information processing of preschool children. *Contributions to Music Education, 7,* 2-11.

McKernon, P. (1979). The development of first songs in young children. *New Directions for Child Development, 3,* 43-58.

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist, 35,* 1012-1027.

Messick, S. (1983). The past and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 33-45). Hillsdale, NJ: Lawrence Erlbaum.

Messick, S. (1989a). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*(2), 5-11.

Messick, S. (1989b). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.

Millman, J. & Greene, J. (1989). The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 335-366). New York: Macmillan.

Monahan, C. B., & Carterette, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception, 3*(1), 1-32.

Monohan, C. B., Kendall, R. A., & Carterette, E. C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. *Perception and Psychophysics, 41,* 576-600.

Moorhead, G. E., & Pond, D. (1978). *Music of young children.* Santa Barbara, CA: Pillsbury Foundation for Advancement of Music Education.

Morrongiello, B. A., Trehub, S. E., Thorpe, L. A., & Capodilupo, S. (1985). Children's perception of melodies: The role of contour, frequency, and rate of presentation. *Journal of Experimental Child Psychology, 40,* 79-292.

Morrongiello, B. A., & Roes, C. L. (1990). Developmental changes in children's perception of musical sequences: Effects of musical training. *Developmental Psychology, 26*(5), 814-820.

Mursell, J. L. (1937). *The Psychology of Music.* New York: W.W. Norton.

Nitko, A. J. (1996). *Educational assessment of students* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Norris, S. P. (1990). Effect of eliciting verbal reports of thinking on critical thinking test performance. *Journal of Educational Measurement, 27* (1), 41-58.

Norusis, M. J. (1993). *SPSS for windows base system user's guide release 6.0.* Chicago: SPSS.

Nye, R. E. & Nye, V. T. (1992). *Music in the elementary school* (6th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Owens, R. G. (1982). Methodological rigor in naturalistic inquiry: Some issues and answers. *Educational Administration Quarterly, 18* (2), 1-21.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 116-126.

Patton, M. Q. (1990). *Qualitative evaluation and research methods.* Newbury Park, CA: Sage.

Pederson, D. M., & Pederson, N. O. (1970). The relationship between pitch recognition and vocal pitch production in sixth-grade students. *Journal of Research in Music Education, 18*(3), 265-272.

Petzold, R. G. (1960). The perception of music symbols in music reading by normal children and by children gifted musically. *Journal of Experimental Education, 28*(4), 271-319.

Petzold, R. G. (1966). *Auditory perception of musical sounds by children in the first six grades.* Madison: University of Wisconsin. (ERIC Document Reproduction Service No. ED 010 297).

Piston, W., & DeVoto, M. (1987). *Harmony* (5th ed.). New York: W.W. Norton.

Radocy, R. (1989). Evaluating student achievement. *Music Educators Journal, 76*(4), 30-33.

Radocy, R. E., & Boyle, J. D. (1988). *Psychological foundations of musical behavior* (2nd ed.). Springfield, IL: Charles C. Thomas.

Rainbow, E. L. (1965). A pilot study to investigate the constructs of musical aptitude. *Journal of Research in Music Education, 13*(1), 3-14.

Ramsey, J. H. (1983). The effects of age, singing ability and instrumental experiences on preschool children's melodic perception. *Journal of Research in Music Education, 31*(2), 133-145.

Reimer, B. (1989). *A philosophy of music education* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Revesz, G. (1953). *Introduction to the psychology of music.* London: Longmans.

Rogers, G. L. (1991). Effect of color-coded notation on music achievement of elementary instrumental students. *Journal of Research in Music Education, 39*(1), 64-73.

Rowntree, J. P. (1970). The Bentley measures of musical abilities: A critical evaluation. *Bulletin of the Council For Research in Music Education, 22,* 25-32.

Scardamalia, M. & Bereiter, C. (1983). Child as coinvestigator. In S. G. Paris, G. M. Olson, & H. W. Stevenson (Eds.), *Learning and motivation in the classroom* (pp. 61-82). Hillsdale, NJ: Lawrence Erlbaum.

Scott, C. R. (1979). Pitch concept formation in preschool children. *Bulletin of the Council For Research in Music Education, 59,* 87-93.

Seashore, C. E. (1919). *The psychology of musical talent.* Boston: Silver Burdett.

Seashore, C.E. (1938). *Psychology of Music.* New York: McGraw-Hill.

Seashore, C. E. (1947). *In search of beauty in music.* New York: McGraw-Hill.

Seashore, C. E., Lewis, D., & Saetveit, J. G. (1960). *Seashore measures of musical talents* (rev. ed.). New York: The Psychological Corporation.

Serafine, M. L. (1988). *Music as cognition: The development of thought in sound.* New York: Columbia University.

Sergeant, D. (1973). Measurement of pitch discrimination. *Journal of Research in Music Education, 21*(1), 3-19.

Sergeant, D., & Roche, S. (1973). Perceptual shifts in the auditory information processing of young children. *Psychology of Music, 1*(2), 39-48.

Shepard, R. N. (1981). Individual differences in the perception of musical pitch. In *Documentary report of the Ann Arbor symposium: National symposium on the applications of psychology to the teaching and learning of music* (pp. 152-174). Reston, VA: Music Educators National Conference.

Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. *Psychology Review, 89,* 305-333.

Shuter-Dyson, R., & Gabriel, C. (1981). *The psychology of musical ability* (2nd ed.). London: Methuen.

Simons, G. (1986). Early childhood musical development: A survey of selected research. *Bulletin of the Council For Research in Music Education, 86,* 36-52.

Sloboda, J. A. (1974). The eye-hand approach to the study of sight reading. *Psychology of Music, 2*(2), 237-244.

Sloboda, J.A. (1976a). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology, 28,* 1-16.

Sloboda, J. A. (1976b). The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Canadian Journal of Psychology, 30,* 228-236.

Sloboda, J. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology, 68,* 117-24.

Sloboda, J. (1978). Perception of contour in music reading. *Perception, 7,* 323-331.

Sloboda, J.A. (1985). *The musical mind: The cognitive psychology of music.* Oxford: Clarendon Press.

Small, R. A., & McCachern, F. L. (1983). The effect of male and female vocal modeling on pitch-matching accuracy of first-grade children. *Journal of Research in Music Education, 31*(3), 227-234.

Staton, B., Staton, M., Davidson, M., Kaplan, M., & Snyder, S. (1988). *Music and You (Grade 4): Teacher's Edition.* New York: Macmillan.

Staton, B., Staton, M., Davidson, M., & Ferguson, N. (1988). *Music and You (Grade 5): Teacher's Edition.* New York: Macmillan.

Staton, B., Staton, M., Davidson, M., Davis, A., & Ferguson, N. (1988). *Music and You (Grade 6): Teacher's Edition.* . New York: Macmillan.

Steen, A. (1992). *Exploring Orff: A teacher's guide.* New York: Schott Music.

Taylor, J. A. (1976). Perception of tonality in short melodies. *Journal of Research in Music Education, 24*(4), 197-208.

Tenopyr, M. L. (1977). Content-construct confusion. *Personnel Psychology, 30,* 47-54.

Tohidi, N. E., Steinkamp, M. W., & Maehr, M. L. (1986). *Gender differences in performance on tests of cognitive functioning: A meta-analysis of research findings.* (ERIC Document Reproduction Service No. 302 452).

Trehub, S. E., Cohen, A. J., Thorpe, L. A., & Morrongiello, B. A. (1986). Development of the perception of musical relations: Semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance, 12*(3), 295-301.

Van Zee, N. (1976). Responses of kindergarten children to musical stimuli and terminology. *Journal of Research in Music Education, 24*(1), 14-21.

Webster, P. R., & Schlentrich, K. (1982). Discrimination of pitch direction by preschool children with verbal and nonverbal tasks. *Journal of Research in Music Education, 30*(3), 151-162.

Whybrew, W. E. (1971). *Measurement and evaluation in music* (2nd ed.). Dubuque, IA: Wm. C. Brown.

Willett, B. E., & Netusil, A. J. (1989). Music computer drill and learning styles at the fourth-grade level. *Journal of Research in Music Education, 57*(3), 219-229.

Wing, H. D. (1941). A factorial study of musical tests. *British Journal of Psychology, 31,* 341-355.

Wing, H. D. (1968). *Tests of musical ability and appreciation: An investigation into the measurement, distribution, and development of musical capacity* (2nd ed.). Cambridge: University Press.

Yarbrough, C. (1991). Inaccurate singers: An exploratory study of variables affecting pitch-matching. *Bulletin of the Council For Research in Music Education, 107,* 23-34.

Young, W. T. (1976). A longitudinal comparison of four music achievement and four music aptitude tests. *Journal of Research in Music Education, 24*(3), 97-109.

Zdzinski, S. F. (1992). Relationships among parental involvement, music aptitude, and musical achievement of instrumental music students. *Journal of Research in Music Education, 40*(2), 114-125.

Zerull, D. S. (1990). Evaluation in arts education: Building and using an effective assessment strategy. *Design for Arts Education, 92*(1), 19-24.

Zimmerman, M. P. (1985). State of the art in early childhood music and research. In J. Boswell (Ed.), *The young child and music: Contemporary principles in child development and music education: Proceedings of the music in early childhood conference* (pp. 65-78). Reston, VA: Music Educators' National Conference.

APPENDIX A

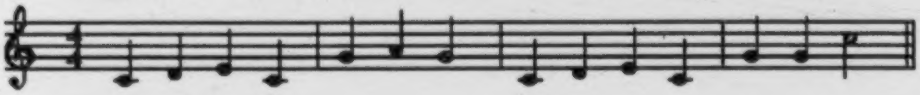
INITIAL TEST ITEMS

Music Reading Test

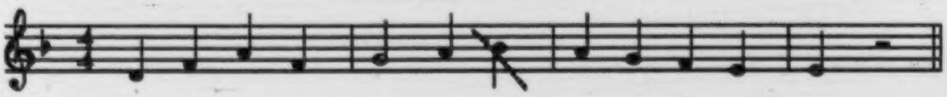
This is a music reading test. All questions are about melody. All directions and questions for this test are on the cassette tape. Once the test has started, you must stay with the tape at all times. If you do not have time to complete a question, make a guess and go on to the next one. If your pencil breaks, put up your hand and you will be given another one.

Before the test begins, we will complete the following practise question together:

The first part of the melody written below will be sung one time. Follow the music notation and put a line through the last note of the melody that you hear sung aloud.



The answer to this question is shown below:



Your teacher will now stop the tape. Ask any questions you have about the directions.

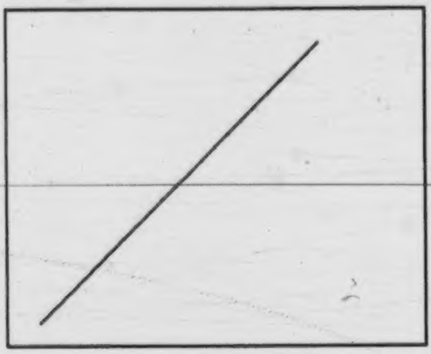
Remember: **THE TAPE CANNOT BE STOPPED ONCE THE TEST BEGINS.**

Turn the page to question 1.

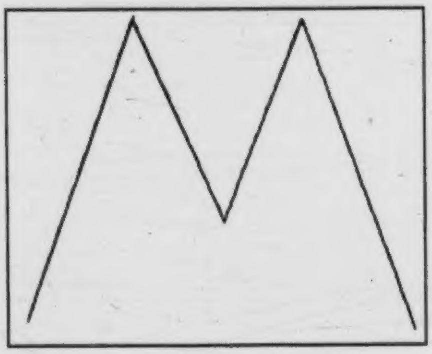
1. You will hear a melody sung two times. Choose the one picture which best describes the overall shape or contour of the melody. Write your answer on the blank provided.

____

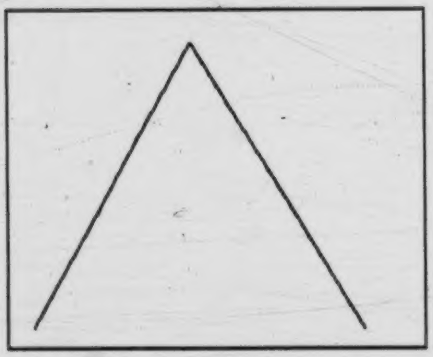
A.



B



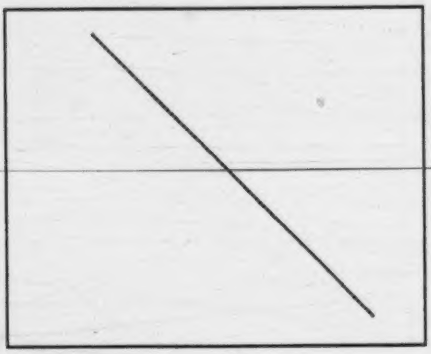
C.



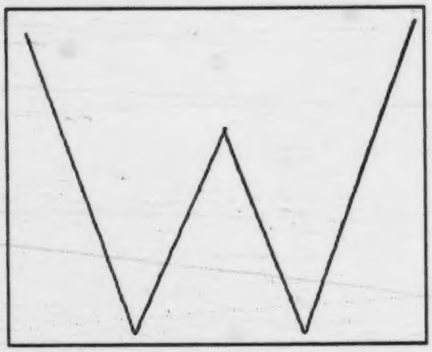
2. You will hear a melody sung two times. Choose the one picture which best describes the overall shape or contour of the melody. Write your answer on the blank provided.

____

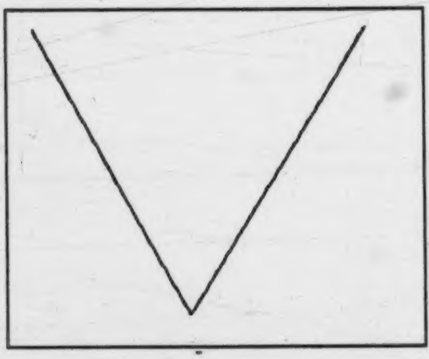
A.



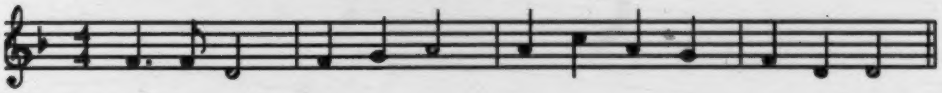
B.



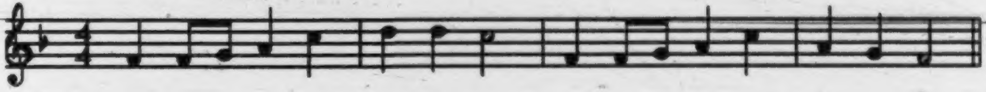
C.



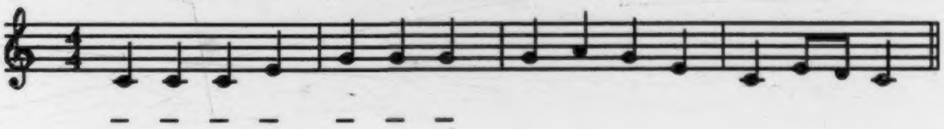
3. The melody written below will be sung one time. Follow the music notation and circle all the notes (♩♩ ♩♩♩) that repeat.



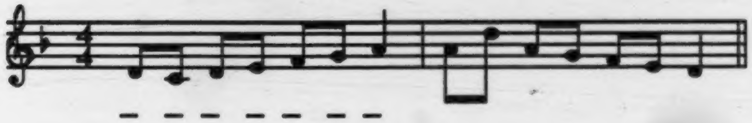
4. The melody written below will be sung one time. Follow the music notation and circle all the notes (♩♩ ♩♩♩) that move by step.



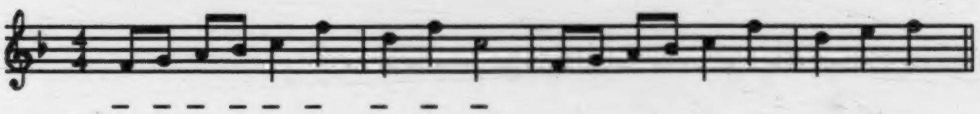
5. You will hear the melody written below sung one time. Write the correct solfa syllables (d,r,m,f,s,l,t) on the lines underneath the notes.



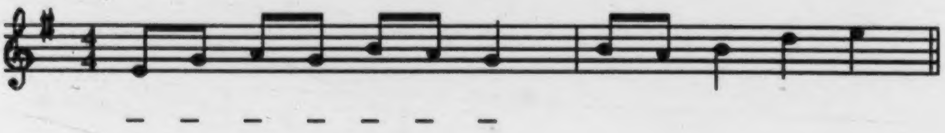
6. You will hear the melody written below sung one time. Write the correct solfa syllables (d,r,m,f,s,l,t) on the lines underneath the notes.



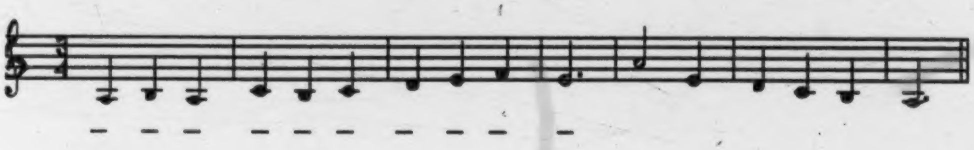
7. You will hear the melody written below sung one time. Write the absolute letter names (A,B,C,D,E,F,G) on the lines underneath the notes.



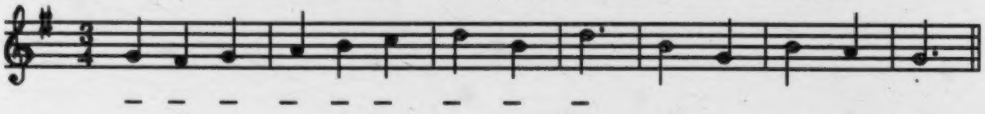
8. You will hear the melody written below sung one time. Write the absolute letter names (A,B,C,D,E,F,G) on the lines underneath the notes.



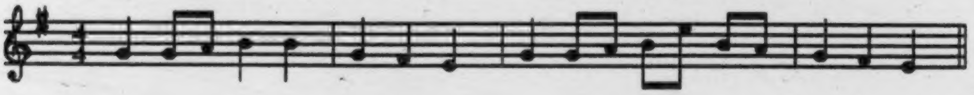
9. You will hear the melody written below sung one time. Write the correct solfa syllables (d,r,m,f,s,l,t) or the absolute letter names (A,B,C,D,E,F,G) on the lines under the notes. Then circle the notes which are separated by a half step (minor second).



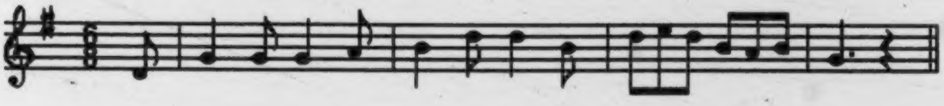
10. You will hear the melody written below sung one time. Write the correct solfa syllables (d,r,m,f,s,l,t) or absolute letter names (A,B,C,D,E,F,G) on the lines under the notes. Then circle the notes which are separated by a half step (minor second).



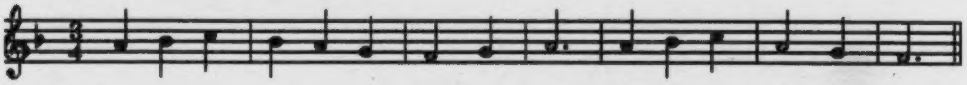
11. The first part of the melody written below will be sung one time. Follow the music notation and put a line through the last note of the melody that you hear sung aloud.



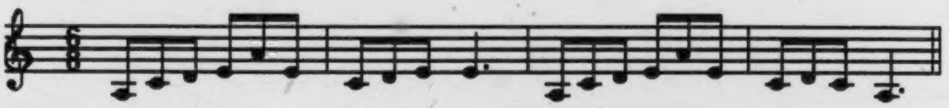
12. The first part of the melody written below will be sung one time. Follow the music notation and put a line through the last note of the melody that you hear sung aloud.



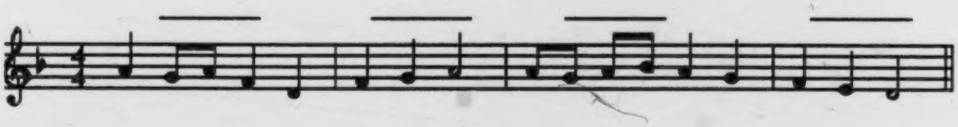
13. Only part of the melody written below is sung aloud. In order to follow along, you must hear the melody in your head. As you follow the notation, put a line through the notes that you do not hear sung aloud.



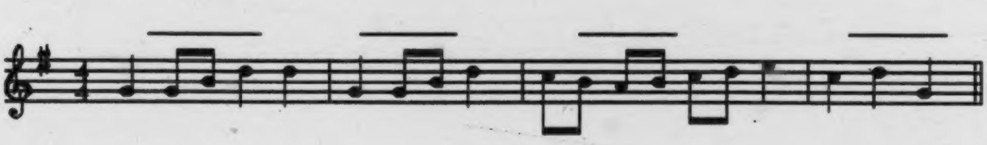
14. Only part of the melody written below is sung aloud. In order to follow along, you must hear the melody in your head. As you follow the notation, put a line through the notes that you do not hear sung aloud.



15. You will hear a melody sung two times. Place a check mark (√) on the line above the one measure in which the music notation is different than the melody you hear.



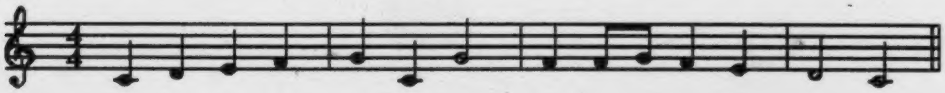
16. You will hear a melody sung two times. Place a check mark (√) on the line above the one measure in which the music notation is different than the melody you hear.



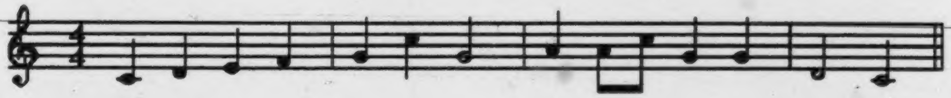
17. You will hear a melody sung two times. Choose the one melody which is the same as the melody you hear. Write your answer on the blank provided.

____

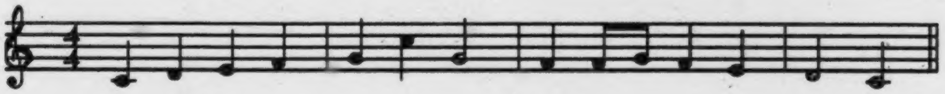
A.



B.



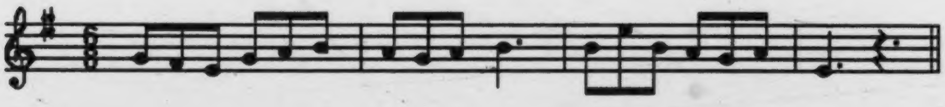
C.



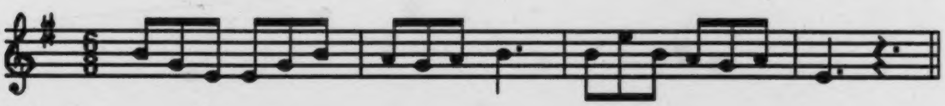
18. You will hear a melody sung two times. Choose the one melody which is the same as the melody you hear. Write your answer on the blank provided.

____

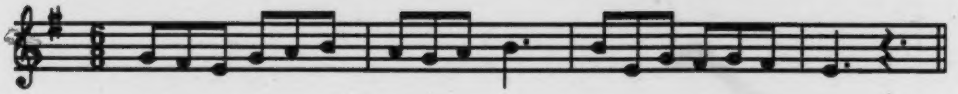
A.



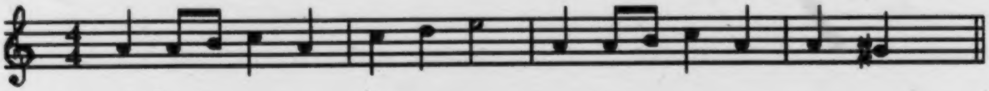
B.



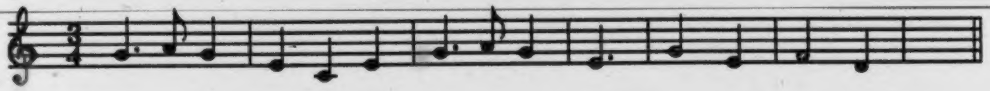
C.



19. The melody written below will be sung one time. The last note is missing. Complete this melody so that it ends on the home tone or tonal center.



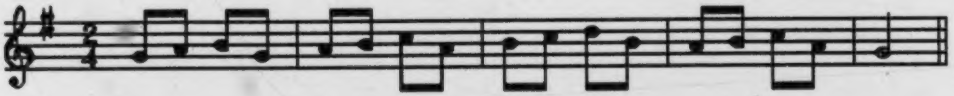
20. The melody written below will be sung one time. The last note is missing. Complete this melody so that it ends on the home tone or tonal center.



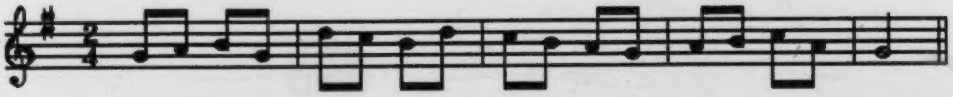
21. You will hear each melody one time. Choose the one melody which best continues the repeating pattern or sequence. Write your answer on the blank provided.

____

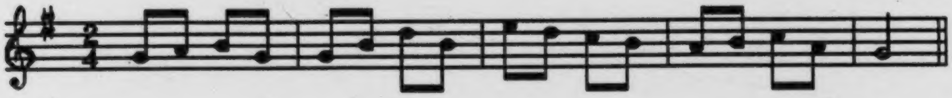
A.



B.



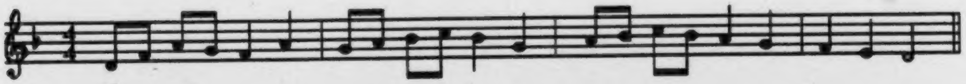
C.



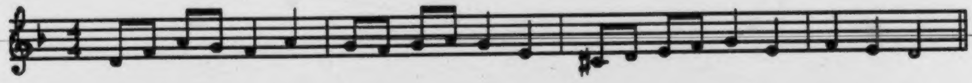
22. You will hear each melody one time. Choose the one melody which best continues the repeating pattern or sequence. Write your answer on the blank provided.

____

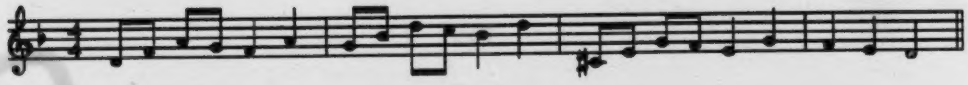
A.



B.



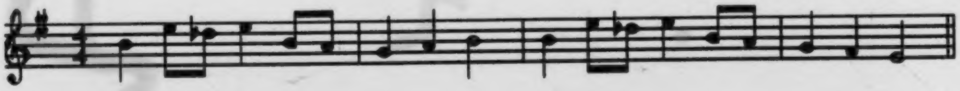
C.



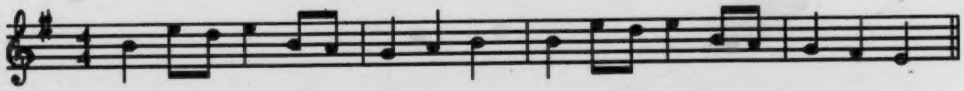
23. You will hear a pair of melodies sung two times. Listen for the sharps or flats. Choose the notation which is the same as the second melody you hear. Write your answer on the blank provided.

____

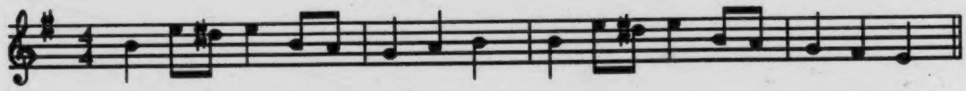
A.



B.



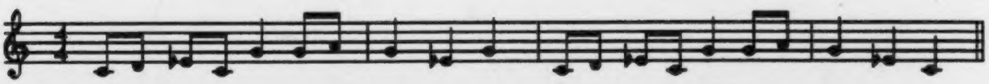
C.



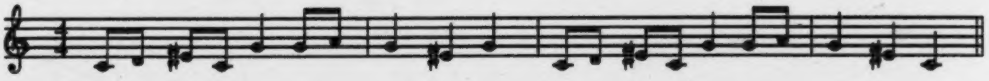
24. You will hear a pair of melodies sung two times. Listen for the sharps or flats. Choose the notation which is the same as the second melody you hear. Write your answer on the blank provided.

____

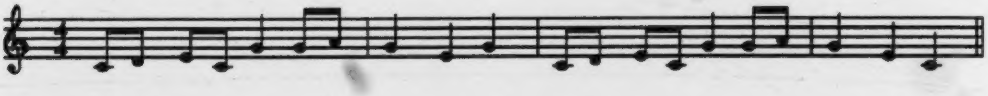
A.



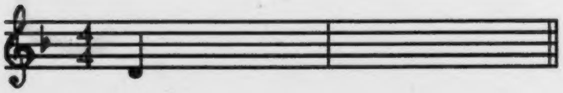
B.



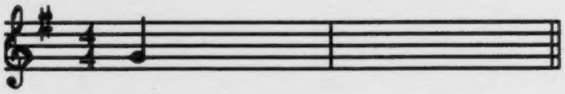
C.



25. You will hear a melody sung three times. Write the melody on the staff below. The first note is given.



26. You will hear a melody sung three times. Write the melody on the staff below. The first note is given.



APPENDIX B

QUESTIONNAIRE COMPLETED BY SUBJECT-AREA SPECIALISTS: ILLUSTRATION OF STRUCTURE

Questionnaire

The content of this music reading test is based upon a survey of selected textbook series and methodologies in which music literacy is viewed, at least in part, as the ability to read and/or write music notation. In these sources melodic perception is subdivided into several different characteristics. In this test, these various characteristics of melody are assessed in thirteen Test Clusters.

Each test cluster begins with an characteristic of melody identified in the survey described above. This is followed by a general description which illustrates how the characteristic of melody is related to the corresponding test items. Each cluster is comprised of two items (for example, cluster #2 contains items 2A and 2B). For each item, please answer each of the following four questions:

1. Reading Level:
   Is the reading level of the item appropriate for students in grade 5?

2. Irrelevant Material:
   Is the item free of irrelevant material?

3. Representativeness:
   This test assesses music reading skills within a group situation. With this in mind, how well do you think the general description and sample test items assess music achievement in terms of the characteristic of melody?

4. Relevance:
   How well do you think the characteristic of melody, illustrated in the sample items, relates to melodic perception as it is portrayed in elementary school music programs?

Test Cluster #2

1. **Characteristic of Melody**- Melodic motion consists of notes which move by step or skip. The same note may be repeated in succession.

2. **General Description**- Students listen to a previously unheard melody (presented one time) and, depending upon the particular item, identify points at which the melody moves by step, skip, or repeats.

The following four questions refer to test items 2A and 2B, reproduced on the following page.

Respond to questions 1 and 2 with √ for 'yes', X for 'no', and '?' for unsure. Please write your response in the appropriate box below.

1. Is the reading level of the item appropriate for students in grade 5?

Cluster	2	2
Item	2A	2B
Response		

If 'no', please write suggestions for improvement on the test items.

2. Is the item free of irrelevant material?

Cluster	2	2
Item	2A	2B
Response		

If 'no', please write suggestions for improvement on the test item.

Respond to questions 3 and 4 in terms of the following rating scale: 1- poor, 2- fair, 3- good, 4- very good, 5- excellent. Please write the number corresponding to your rating in the appropriate box below.

3. This test assesses music reading skills within a group situation. With this in mind, how well do you think the general description and sample test items assess music achievement in terms of the characteristic of melody.

Cluster	2	2
Item	2A	2B
Response		

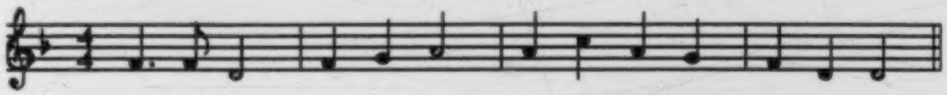
4. How well do you think the characteristic of melody **'melodic motions consists of notes which move by step or skip,'** illustrated in the sample items, relates to melodic perception as it is portrayed in school music programs?

Cluster	2	2
Item	2A	2B
Response		

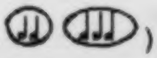
Please write any additional suggestions for improvement on the test item.

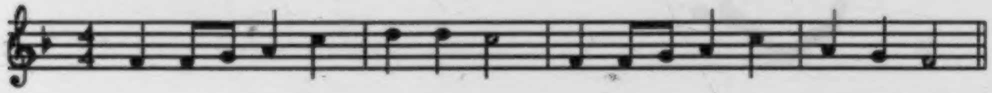
Item 2A (test item 3)

The melody written below will be sung one time. Follow the music notation and circle all the notes ( ) that repeat.



Item 2B (test item 4)

The melody written below will be sung one time. Follow the music notation and circle all the notes () that move by step.



Note: Using the questionnaire structure illustrated here for items in test cluster 2, music educators assessed all 26 items items developed in this study.

APPENDIX C

LETTERS TO PARTICIPANTS

Sheila J. Scott
Graduate Student
Department of Elementary Education
Faculty of Education
University of Alberta
T6G 2G5

Dear Parent or Guardian,

I am a PhD student in Elementary Education at the University of Alberta and am developing a music reading test for use in the upper elementary grades. The test is divided into two parts. In the first part students will hear melodies sung by a soprano voice. They are instructed to answer questions corresponding to what they hear. In the second part they will answer similar questions without hearing the melody. The test will be completed in a regularly scheduled music class.

As part of this study, I will be administering the test to grade 5 students at --------------------- School. I am writing to request permission for your child to take part in this study. Participation is voluntary and you may withdraw at any time. Test results will be used in a detailed analysis of test items. Participants' names will not be used in the final report and the tests will be destroyed after the completion of the study. If you have further questions, please contact Sheila Scott at -----------.

Please complete the bottom portion of this letter and return it to your child's music teacher, Mrs. -----------------, by ---------------.

Sincerely,

Sheila Scott	----------------
Graduate Student	Principal
The University of Alberta	------------------

Please complete the following and return to Mrs. ----- by -------.
Child's Name- ____________

______ - I will allow my child to participate in the study.
______ - I will not allow my child to participate in this study.

____________________	______________________
Date	Signature

Sheila J. Scott
Graduate Student
Department of Elementary Education
Faculty of Education
University of Alberta
T6G 2G5

Dear Parent or Guardian,

I am a PhD student in Elementary Education at the University of Alberta and am developing a music achievement test for use in the upper elementary grades. As part of this study, I will be administering test questions to select grade five students at ----------------- School on an individual basis. I am writing to request permission for your child to take part in this study. In this administration, your child will explain how test questions are answered by 'thinking aloud' while completing each item. The session will be tape recorded and transcribed. Information supplied in this process will be used to revise individual test items.

This test will take approximately 30 minutes to complete and will be administered during school hours at a time convenient to both your child and the teachers at the school. Participation is voluntary and you may withdraw at any time. Participants' names will not be used in the final report and the tests will be destroyed after the completion of the study. If you have further questions, please contact Sheila Scott at ------------.

Please complete the bottom portion of this letter and return to your child's teacher, Miss. -------, by.

Sincerely,

Sheila Scott
Graduate Student
The University of Alberta

Principal

Please complete the following and return to Miss.------ by --------------.

Child's Name- ____________

______ - I will allow my child to participate in the study.
______ - I will not allow my child to participate in this study.

____________________ Date

______________________ Signature

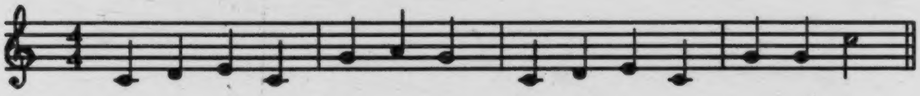
APPENDIX D

TEST VERSION 1

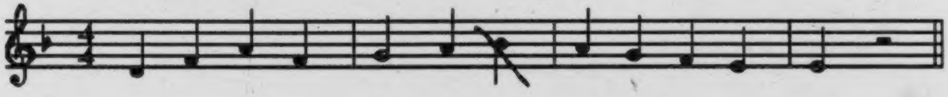
Music Reading Test

This is a music reading test. All questions are about melody. The directions and questions for this test are on the cassette tape. Once the test has started, you must stay with the tape at all times. If you do not have time to complete a question, answer quickly and go on to the next one. If your pencil breaks, put up your hand and you will be given another one. Before the test begins, we will complete the following practise question together:

Listen to the melody once. Follow the music notation and put line through the last note of the melody that you hear sung aloud.



The answer to this question is shown below:



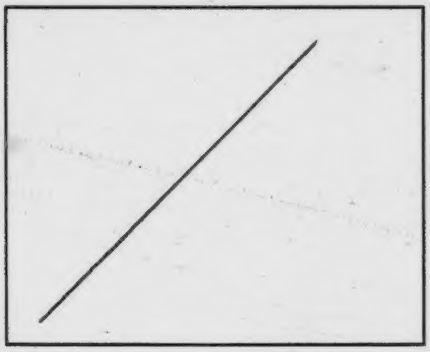
Your teacher will now stop the tape. Ask any questions you have about the directions.

Remember: **THE TAPE CANNOT BE STOPPED ONCE THE TEST BEGINS.**

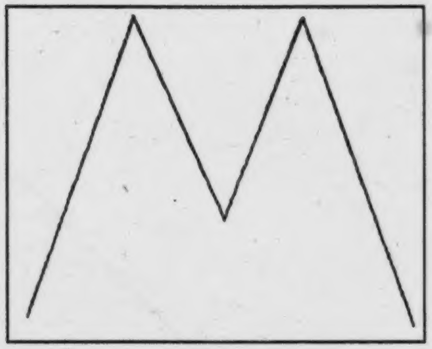
Turn the page to question 1.

1. Listen to the melody twice. Circle the letter above the picture which best describes the **overall shape** or **contour** of the melody.

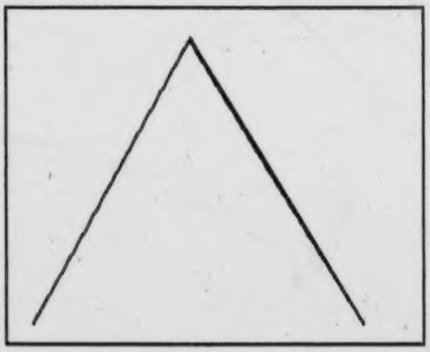
A.



B

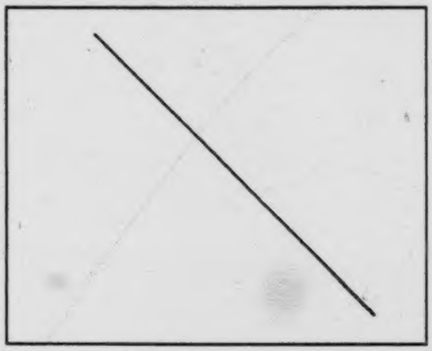


C.

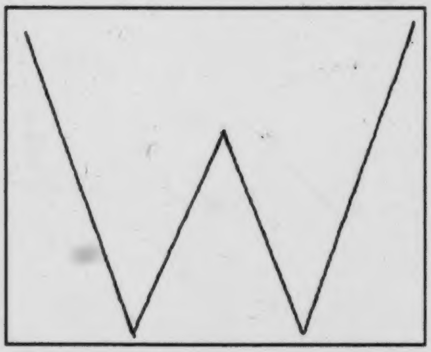


2. Listen to the melody twice. Circle the letter above the picture which best describes the **overall shape** or **contour** of the melody.

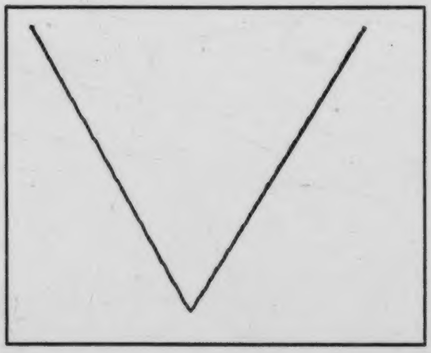
A.



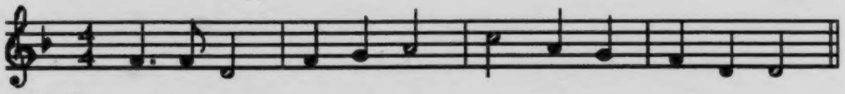
B.



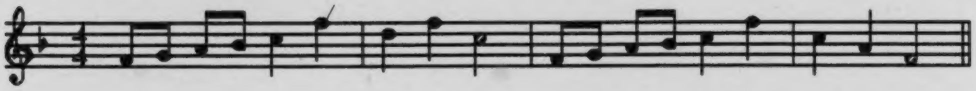
C.



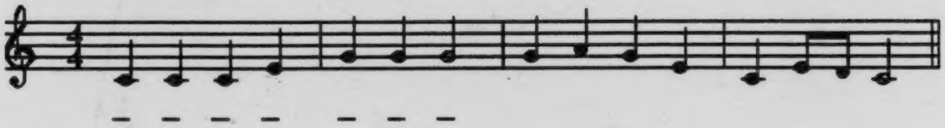
3. Listen to the melody once. Follow the music notation and circle all the groups of notes that **repeat**.



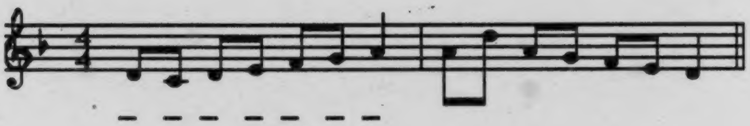
4. Listen to the melody once. Follow the music notation and circle all the groups of notes that move by **step**.



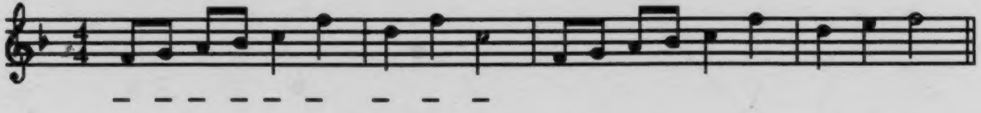
5. Listen to the melody once. Write the **solfa syllables** (d,r,m,f,s,l,t) on the lines under the notes.



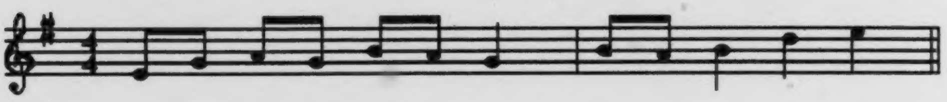
6. Listen to the melody once. Write the **solfa syllables** (l,t,d,r,m,f,s) on the lines under the notes.



7. Listen to the melody once. Write the **absolute letter names** (A,B,C,D,E,F,G) on the lines under the notes.

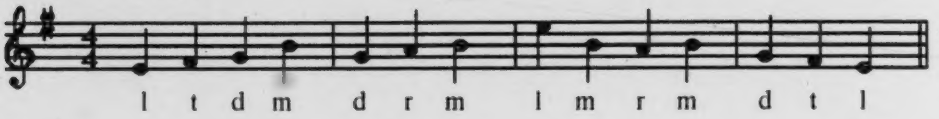


8. Listen to the melody once. Write the **absolute letter names** (A,B,C,D,E,F,G) on the lines under the notes.

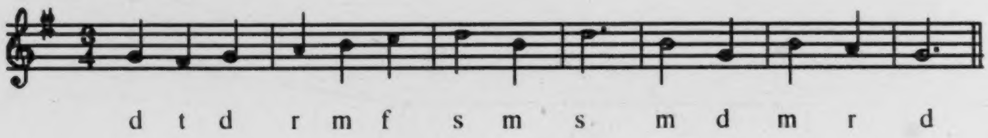


_ _ _ _ _ _ _

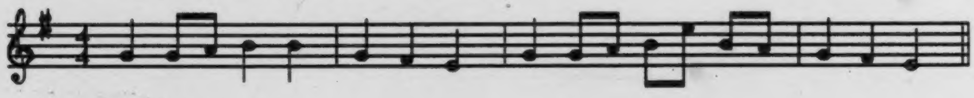
9. Listen to the melody once. Circle the notes which are separated by a **half step (minor second).**



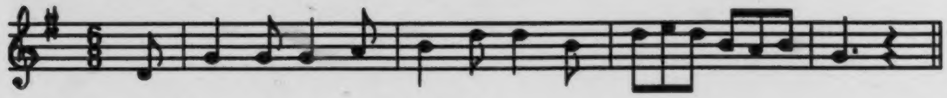
10. Listen to the melody once. Circle the notes which are separated by a **half step (minor second).**



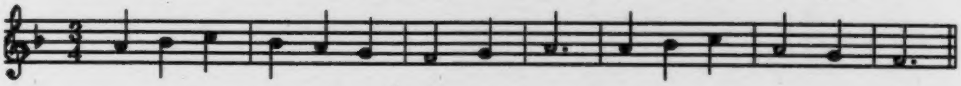
11. Listen to the melody once. Follow the music notation and put a line through the **last note** of the melody that you hear **sung aloud**.



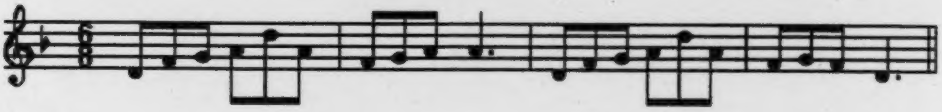
12. Listen to the melody once. Follow the music notation and put a line through the **last note** of the melody that you hear **sung aloud**.



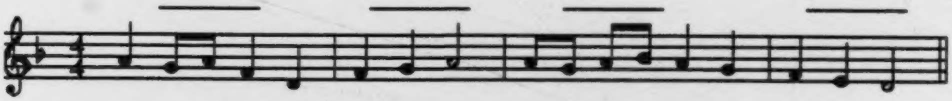
13. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the notation and put a line through the notes that are **not** sung aloud.



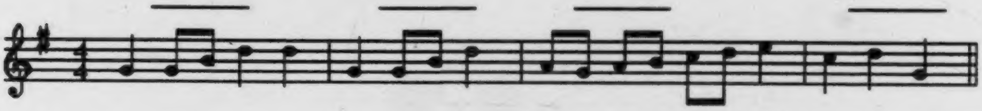
14. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the notation and put a line through the notes that are **not** sung aloud.



15. Listen to the melody twice. Place a check mark (√) above the measure which is **different** than the melody you hear.

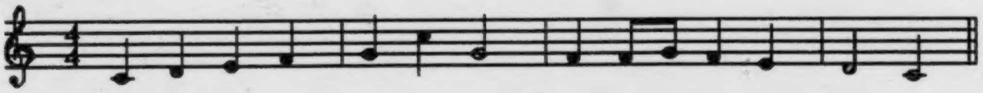


16. Listen to the melody twice. Place a check mark (√) above the measure which is **different** than the melody you hear.

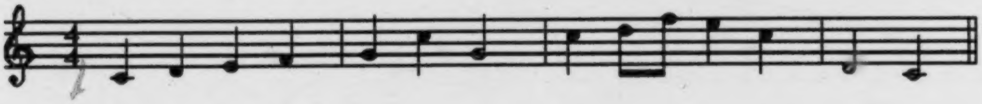


17. Listen to the melody twice. Circle the letter above the melody which is the **same** as the melody you hear.

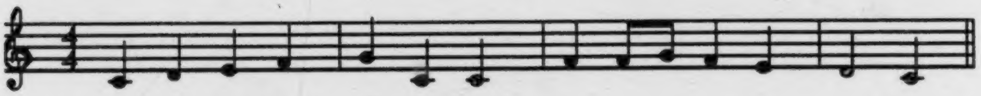
A.



B.

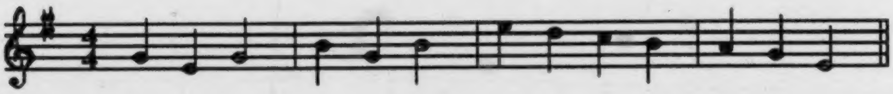


C.

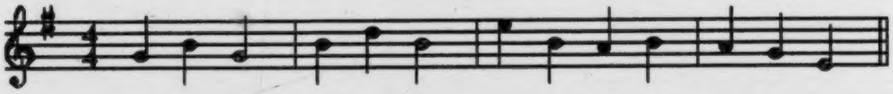


18. Listen to the melody twice. Circle the letter above the melody which is the **same** as the melody you hear.

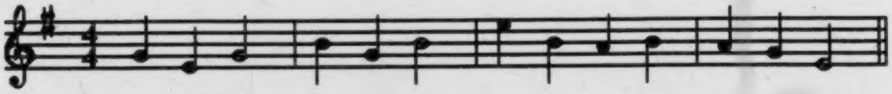
A.



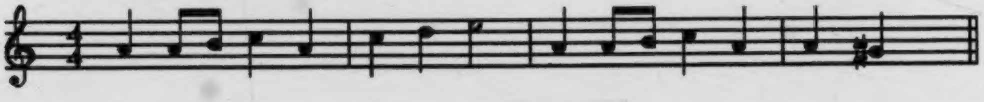
B.



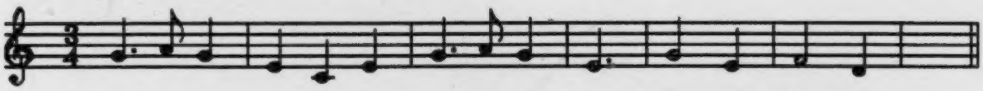
C.



19. Listen to the melody twice. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**.

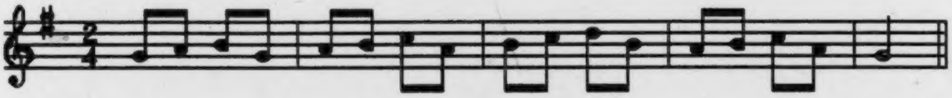


20. Listen to the melody twice. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**.

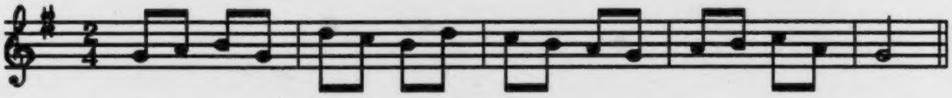


21. Listen to each melody once. Circle the letter above the melody which **best continues** the **repeating pattern** or **sequence**.

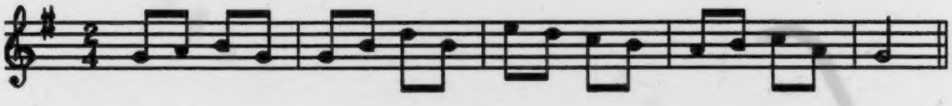
A.



B.

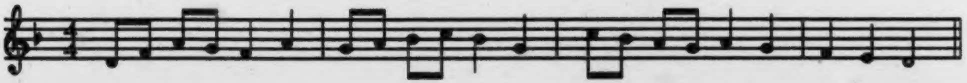


C.

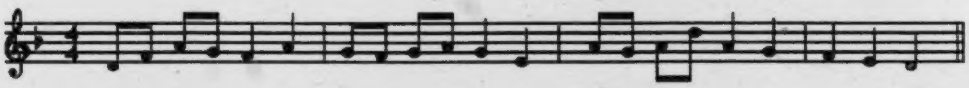


22. Listen to each melody once. Circle the letter above the melody which **best continues** the **repeating pattern** or **sequence**.

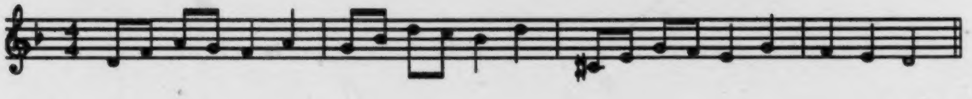
A.



B.

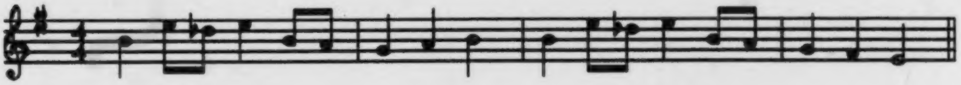


C.

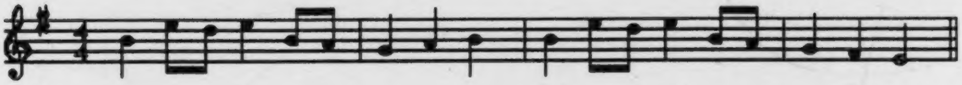


23. Listen to the melody twice. Circle the letter above the melody which is the **same** as the melody you hear.

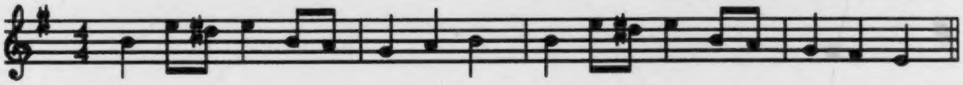
A.



B.

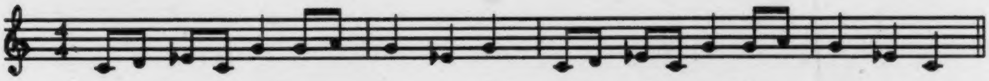


C.

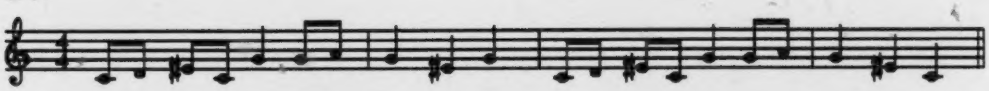


24. Listen to the melody twice. Circle the letter above the melody which is the **same** as the melody you hear.

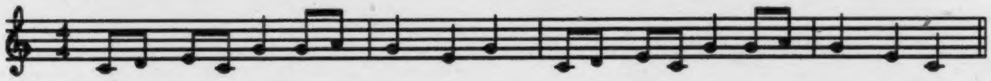
A.



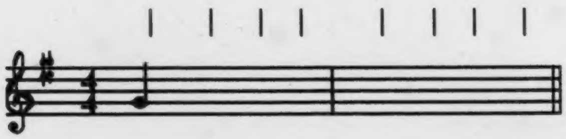
B.



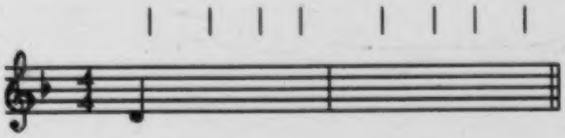
C.



25. Listen to the melody three times. Write the melody on the staff below. The first note and the rhythm are given.



26. Listen to the melody three times. Write the melody on the staff below. The first note and the rhythm are given.



APPENDIX E

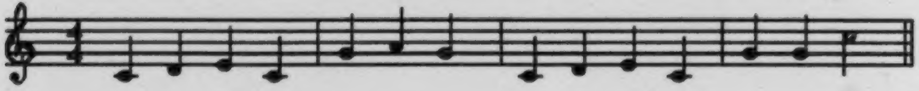
TEST VERSION 2

Music Reading Test

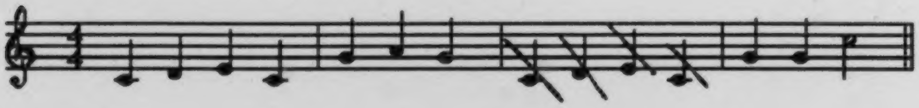
This is a music reading test. All questions are about melody.

Before the test begins, we will complete the following practise question together:

Only part of the melody written below is sung aloud once. As you listen to the melody, follow the notation and put a line through the notes that are **not** sung aloud.



The answer to this question is shown below:



Your teacher will now stop the tape. Ask any questions you have about the directions.

Remember: **THE TAPE CANNOT BE STOPPED ONCE THE TEST BEGINS.**

Turn the page to question 1.

1. DO YOU TAKE PRIVATE MUSIC LESSONS? YES- _____ NO- ___

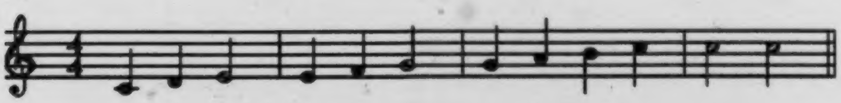
IF ANSWER TO QUESTION 1 IS 'YES':

2. WHAT INSTRUMENT DO YOU PLAY? ______________________

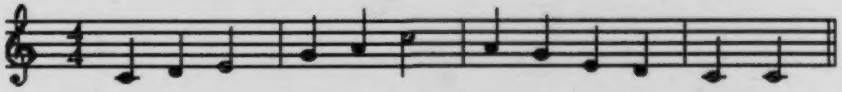
3. HOW MANY YEARS HAVE YOU TAKEN PRIVATE LESSONS? ___

1. Listen to the melody twice. Circle the letter above the melody in which the **overall shape** or **contour** is the same as the melody you hear.

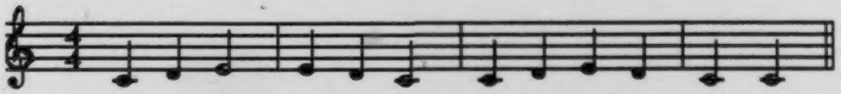
A



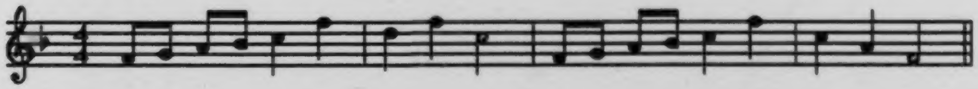
B



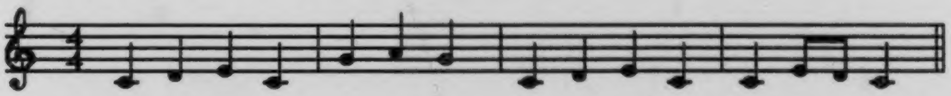
C



2. Listen to the melody once. Follow the music notation and circle all the notes that move by **step**.

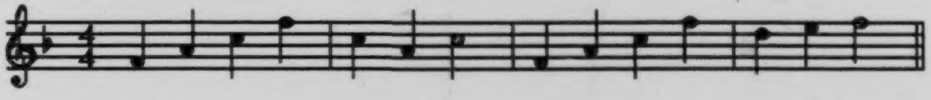


3. Listen to the melody once. Write the **solfa syllables** (d,r,m,f,s,l,t) on the lines under the notes.



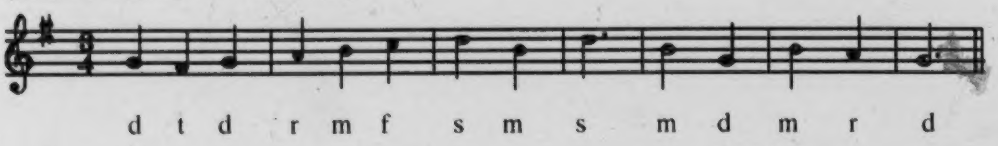
do _ _ _ _ _ _ _

4. Listen to the melody once. Write the **absolute letter names** (A,B,C,D,E,F,G) on the lines under the notes.

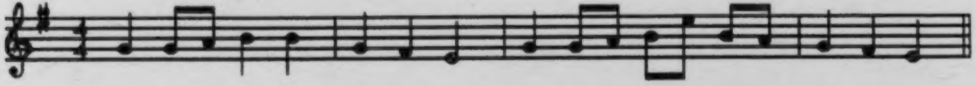


_ _ _ _ _ _ _

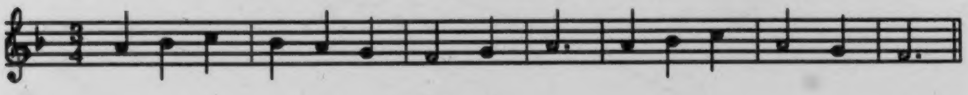
5. Listen to the melody once. Circle the notes which are separated by a **half step (minor second or semitone)**.



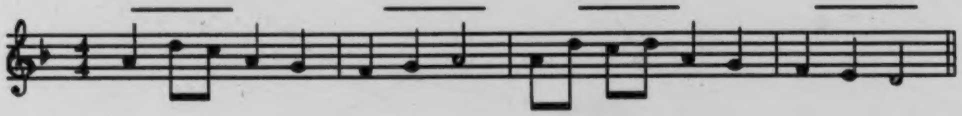
6. Listen to the melody once. Follow the music notation and circle the **last note** of the melody that you hear **sung aloud**.



7. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the notation and put a line through the notes that are **not** sung aloud.

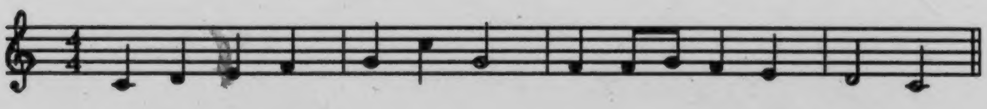


8. Listen to the melody twice. Place a check mark (√) above the **one** measure which is **different** than the melody you hear.

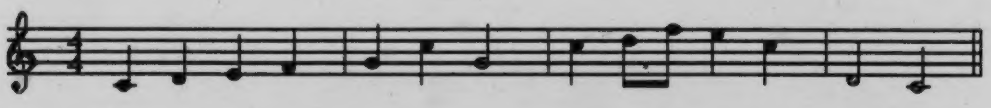


9. Listen to the melody twice. Circle the letter above the one melody which is the **same** as the melody you hear.

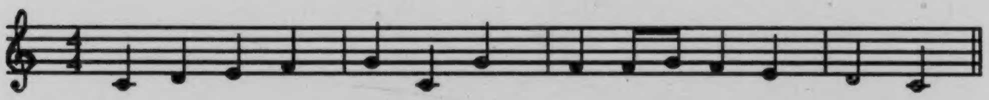
A.



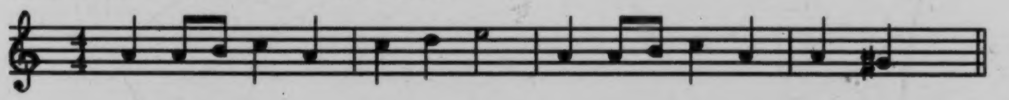
B.



C.

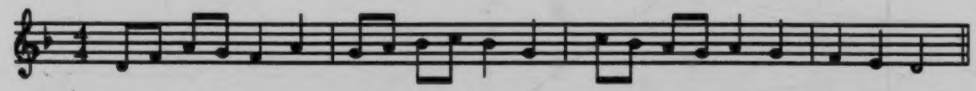


10. Listen to the melody twice. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**.

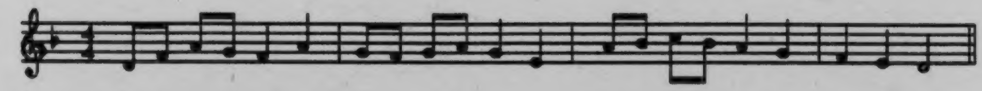


11. Listen to each melody once. The melodies all begin with the same melodic **pattern** or **sequence**. Circle the letter above the music notation which **best continues** the repeating melody.

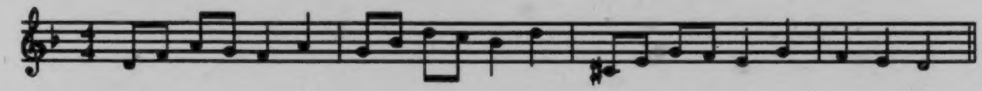
A.



B.

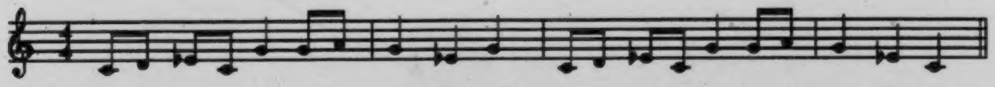


C.

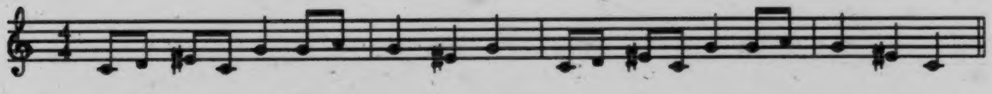


12. Listen to the melody twice. Circle the letter above the one melody which is the **same** as the melody you hear.

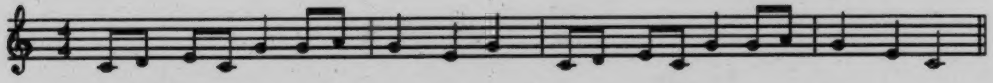
A.



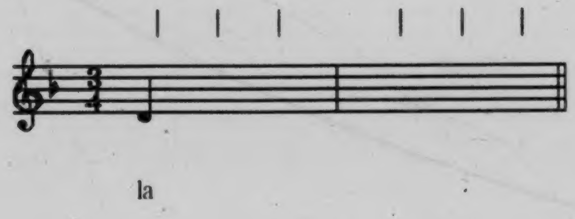
B.



C.



13. Listen to the melody three times. Write the melody on the staff below. The first note and the rhythm are given.



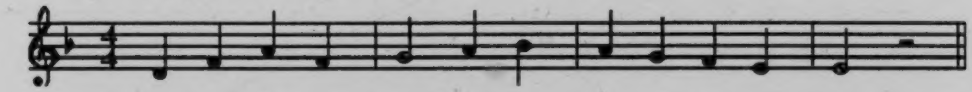
APPENDIX F

TEST VERSION 3

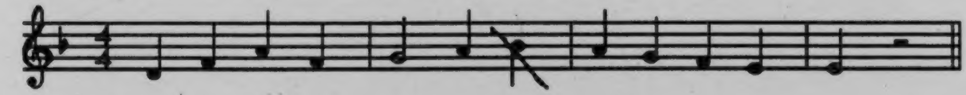
Music Reading Test

This is a music reading test. All questions are about melody. The directions and questions for this test are on the cassette tape. Once the test has started, you must stay with the tape at all times. If you do not have time to complete a question, answer quickly and go on to the next one. If your pencil breaks, put up your hand and you will be given another one.

Before the test begins, we will complete the following practise question together:
Listen to the melody once. Follow the music notation and put a line through the **last note** of the melody you hear sung aloud.



The answer to this question is shown below:



Your teacher will now stop the tape. Ask any questions you have about the directions.

Turn the page to question 1.

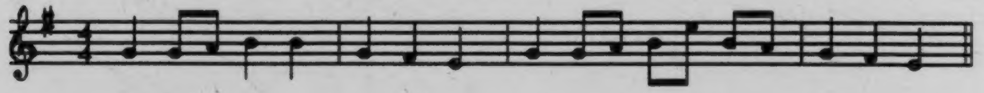
1. DO YOU TAKE PRIVATE MUSIC LESSONS? YES- _____ NO- ____

IF ANSWER TO QUESTION 1 IS 'YES':

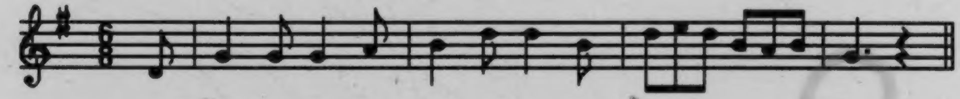
2. WHAT INSTRUMENT DO YOU PLAY? ________________________

3. HOW MANY YEARS HAVE YOU TAKEN PRIVATE LESSONS? ____

1. Listen to the melody once. Follow the music notation and put a line through the **last note** of the melody that you hear **sung aloud**.

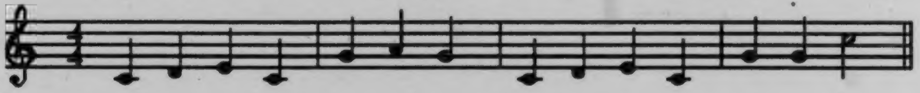


2. Listen to the melody once. Follow the music notation and put a line through the **last note** of the melody that you hear **sung aloud**.

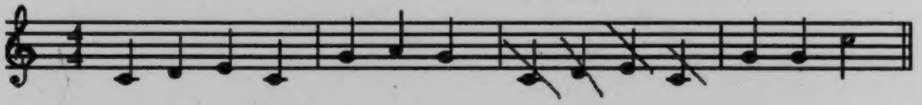


Good! Before we continue we will do another practise question:

Only part of the melody written below is sung aloud once. As you listen to the melody, follow the notation and put a line through the notes that are **not** sung aloud.



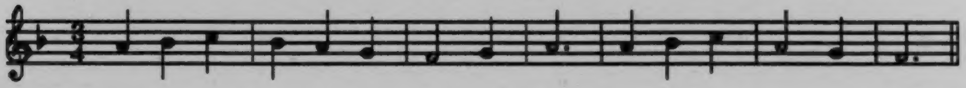
The answer to this question is shown below:



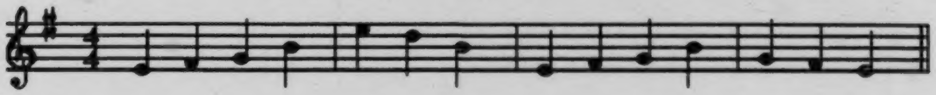
Your teacher will stop the tape. Ask any questions you have about the directions.

Remember: **THE TAPE CANNOT BE STOPPED ONCE THE TEST CONTINUES.**

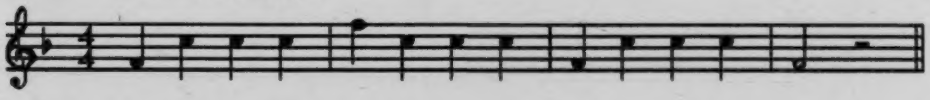
3. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the notation and put a line through the notes that are **not** sung aloud.



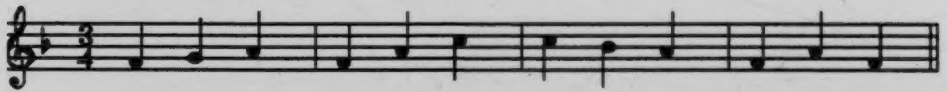
4. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the notation and put a line through the notes that are **not** sung aloud.



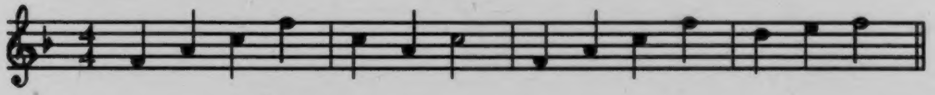
5. Listen to the melody once. Follow the music notation and circle all the notes (♩♩) (♩♩♩) that move by **leap**.



6. Listen to the melody once. Follow the music notation and circle all the notes (♩♩) (♩♩♩) that move by **step**.

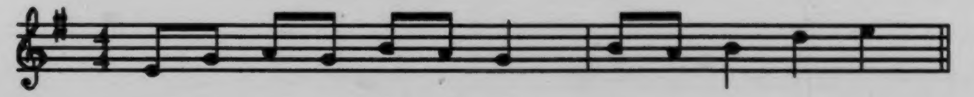


7. Listen to the melody once. Write the **names of the notes** (A,B,C,D,E,F,G) on the lines under the notes.



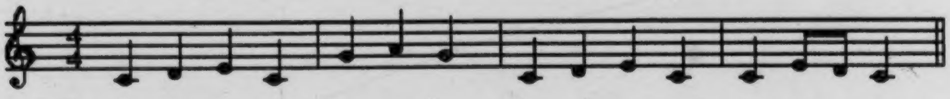
_ _ _ _ _ _ _

8. Listen to the melody once. Write the **names of the notes** (A,B,C,D,E,F,G) on the lines under the notes.



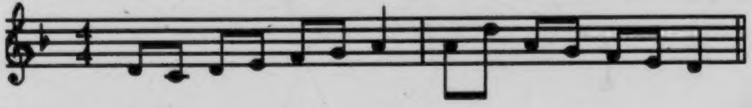
_ _ _ _ _ _ _

9. Listen to the melody once. Write the first letters of the **solfa syllables** (do,re,mi,fa,so,la,ti) on the lines under the notes. The first one is given.



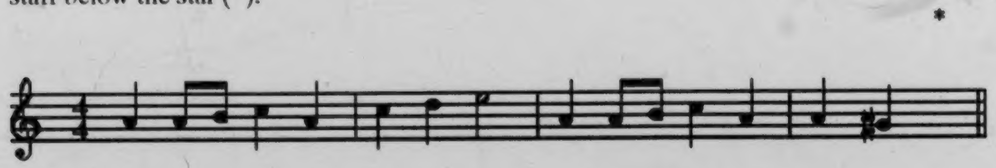
d _ _ _ _ _ _

10. Listen to the melody once. Write the first letters of the **solfa syllables** (la,ti,do,re,mi,fa,so) on the lines under the notes. The first one is given.

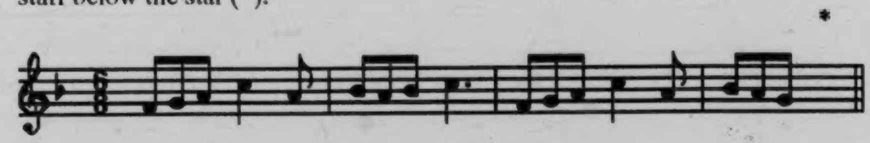


la _ _ _ _ _ _ _

11. Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).

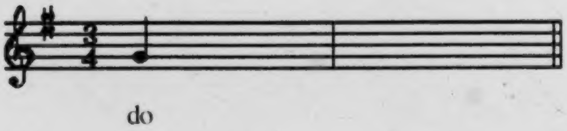


12. Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).



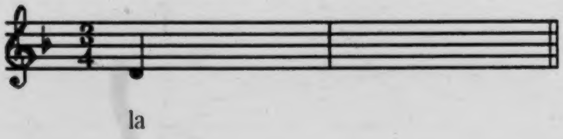
13. Listen to the melody three times. Write the notes of the melody on the staff below. The rhythm is shown above the melody in quarter notes. The first note of the melody is given.

rhythm: ♩ ♩ ♩ ♩ ♩ ♩

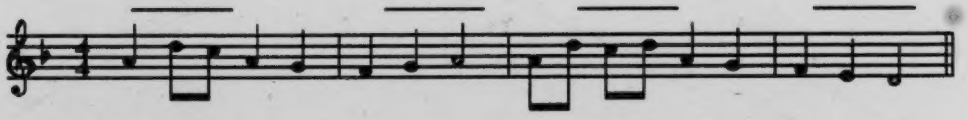


14. Listen to the melody three times. Write the notes of the melody on the staff below. The rhythm is shown above the melody in quarter notes. The first note of the melody is given.

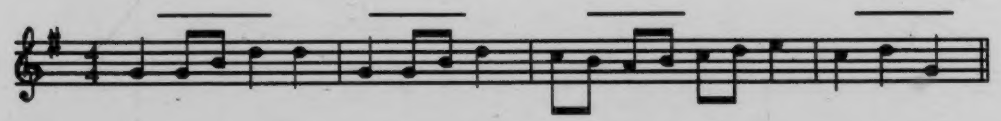
rhythm: ♩ ♩ ♩ ♩ ♩ ♩



15. Listen to the melody twice. Place a check mark (√) above the **one** measure which is **different** than the melody you hear.

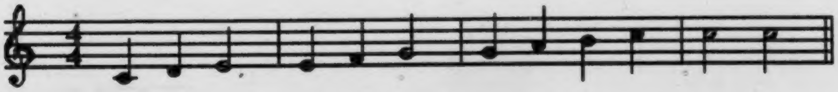


16. Listen to the melody twice. Place a check mark (√) above the **one** measure which is **different** than the melody you hear.

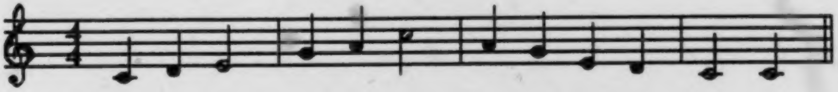


17. Listen to the melody twice. Circle the letter above the **one** melody which is the **same** as the melody you hear.

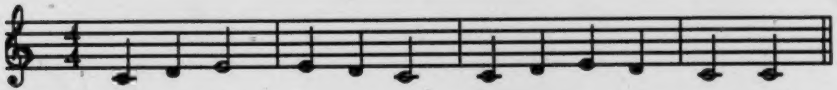
A



B

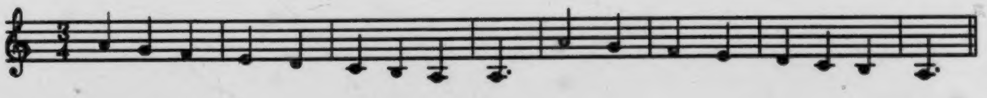


C

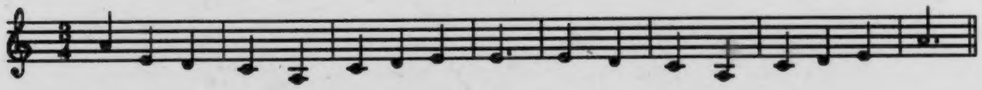


18. Listen to the melody twice. Circle the letter above the **one** melody which it the **same** as the melody you hear.

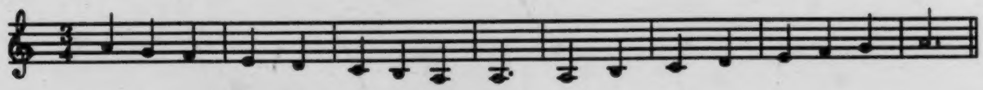
A



B

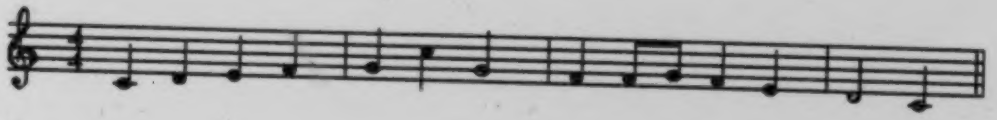


C

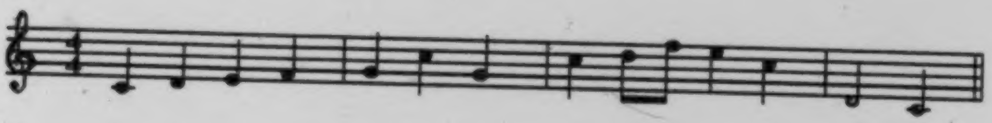


19. Listen to the melody twice. Circle the letter above the __one__ melody which is the **same** as the melody you hear.

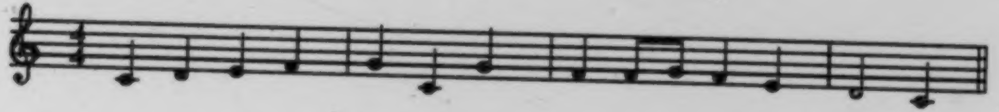
A.



B.

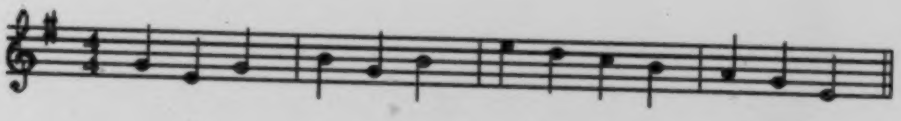


C.

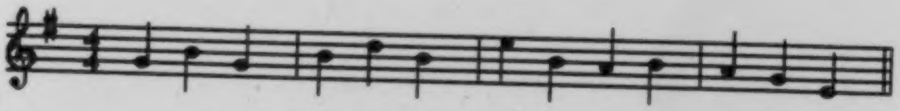


20. Listen to the melody twice. Circle the letter above the __one__ melody which is the **same** as the melody you hear.

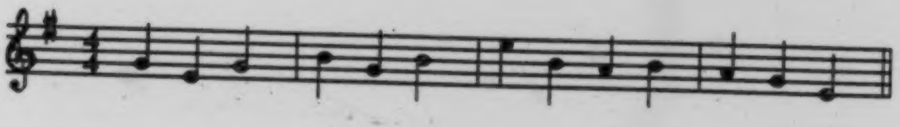
A.



B.

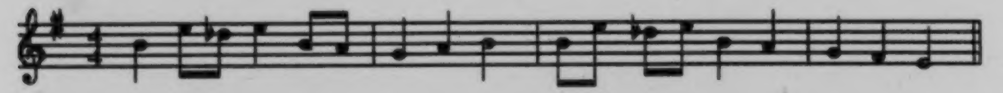


C.

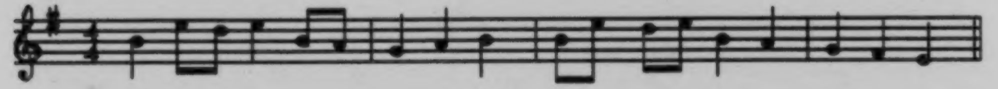


21. Listen to the melody twice. Listen for the sharps or flats and circle the letter above the one melody which is the **same** as the melody you hear..

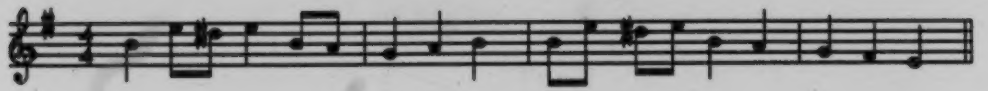
A.



B.

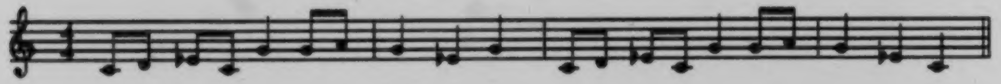


C.

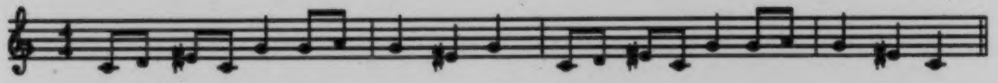


22. Listen to the melody twice. Listen for the sharps or flats and circle the letter above the one melody which is the **same** as the melody you hear.

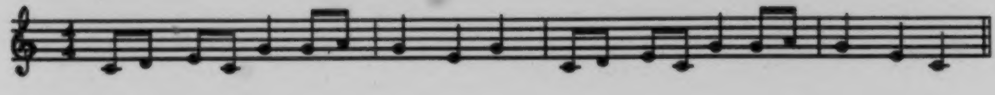
A.



B.

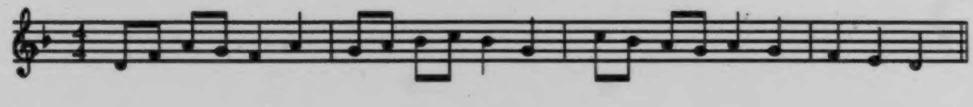


C.

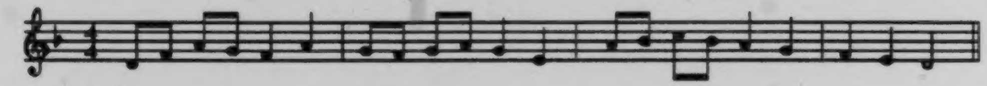


23. Listen to each melody once. The first measure of each melody below begins with the same melodic **pattern** or **sequence**. Measures 2 and 3 are different. Circle the letter above the **one** melody which **best continues** the repeating melody.

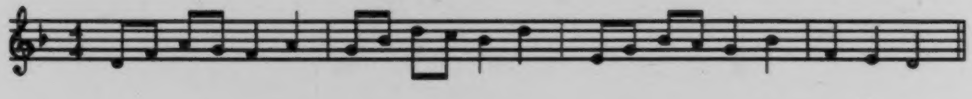
A.



B.

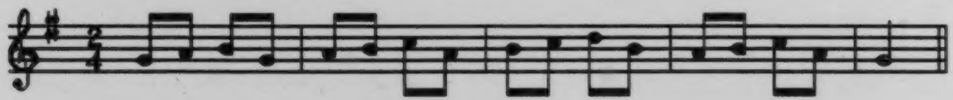


C.

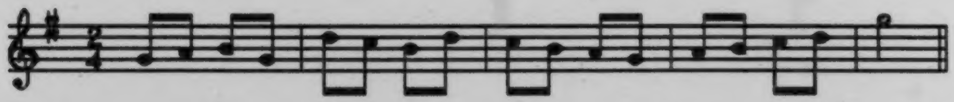


24. Listen to each melody once. The first measure of each melody below begins with the same melodic **pattern** or **sequence**. After that, each melody is different. Circle the letter above the **one** melody which **best continues** the repeating melody.

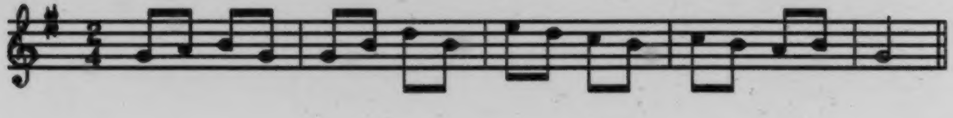
A.



B.



C.



APPENDIX G

FIELD STUDY 3: ITEM-BY-ITEM ANALYSIS FOR ITEMS NOT REVISED

Test Version 3 was administered to students in 4 grade 6 classes in September and October, 1994. All students attended one of 4 schools within School District 2 and each class was taught by one of four music specialists. Of the 24 items administered in Field study 3, 9 were not revised (3, 7, 8, 9, 10, 11, 12, 21, and 22). The item-by-item analysis for each cluster or pair of items is described below. The exception is test item 3, which is discussed in the absence of item 4 (an item which was revised). Unless specified otherwise, item numbers refer to the order in which items were administered in Test Version 3. Numbers in parentheses refer to the order in which the items were administered in Field Study 1.

Item 3 (Item 13 in Field Study 1)

Item 3 was designed to assess a student's ability to follow music notation. In these questions, the aural stimulus was divided into three sections—the first and last were sung aloud and the second was silent. Students were required to follow the notation and indicate which notes were silent.

Indices of Difficulty and Discrimination. The difficulty index for item 3 was .47. The point-biserial index of discrimination was .41.

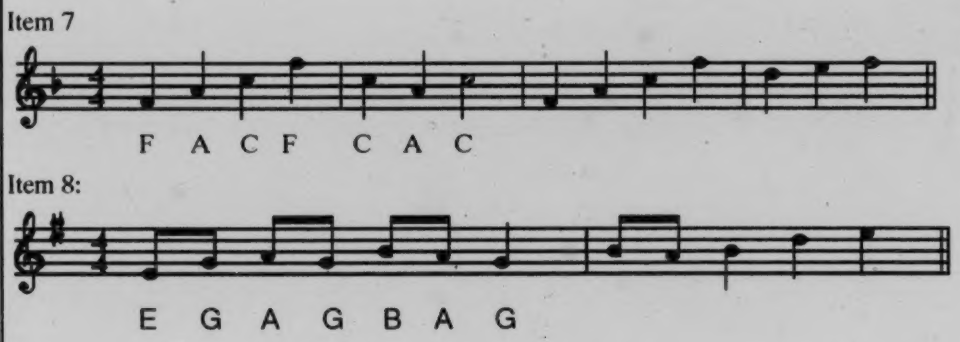
Items 7 and 8 (items 7 and 8 in Field Study 1)

Items 7 and 8 were designed to assess a student's ability to identify notes by their absolute letter names (A, B, C, D, E, F, and G).

Indices of Difficulty and Discrimination. Items 7 and 8 had difficulty indices of .40 and .65, respectively. The point-biserial correlations were .55 and .63, respectively.

As revealed in the data, students more accurately identified the closely-spaced notes in item 8 as opposed to the more widely-spaced notes in item 7. This was surprising from an educational perspective. As illustrated in Field Study 2, in some school music programs note names are taught by equating the spaces of the treble staff to the word face (F-A-C- E) and the lines of the staff to a phrase such as "every good boy does fine" (E-G-B-D-F).

With this in mind, it would seem that item 7, which began with the pattern F-A-C would be easier than item 8 which began with the pattern E-G-A (see Figure 1). Surprisingly, proportionally more students were able to complete item 8, the melody for which was constructed mostly in step-wise motion. This finding seemed to correspond with the research of Goolsby (1989) which demonstrated that music in which the notes are closely spaced is read with fewer and shorter fixations than music in which the notes are spaced further apart.



Note: Students were required to identify the notes in the first two measures of item 7 and the first measure of item 8.

Figure 1. Items 7 and 8: Melodic structure.

Items 9 and 10 (5 and 6)

Items 9 and 10 were designed to assess a student's ability to equate notes with their corresponding solfa syllable names (do, re, mi, fa, so, la, and ti).

Indices of Difficulty and Discrimination. The difficulty index obtained for item 9, .18, was below the accepted guideline of .20. Item 10 had a difficulty index of .45. The point-biserial correlations were .47 and .59, respectively.

In Test Version 3, students were given the solfa syllable corresponding to the first note of the melody (see Field Study 2). In addition, students were given the full names of the solfa syllables rather than an abbreviated form in which each word was represented by its initial syllable. The melody employed for item 9 was revised so that students were required to identify the solfa syllable names of 4 different pitches for which the solfa syllable had not been supplied in the item.

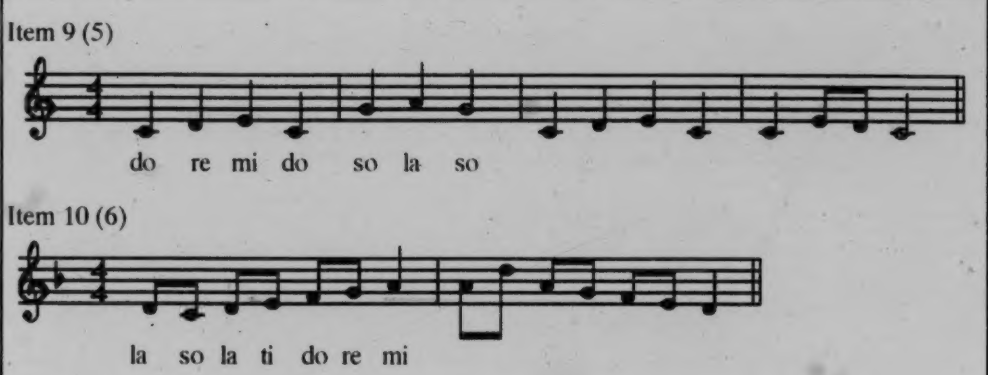
It appears, however, that these revisions influenced performance in only the second of the two items. For item 9, differences in p-level between Field Study 1 ($p = .22$) and Field Study 3 (.18) were negligible, indicating that the same proportion of students were able to write the solfa syllables under the notes regardless of whether the full names of the syllables and the syllable for the first note were given. Considering the number of item revisions, it was not known why this item was so difficult.

It is possible that the relative easiness of item 10 in comparison to item 9 was due to the structure of the melodies. As illustrated in Figure 46, the notes employed in item 9 formed the melody do-re-mi-do-so-la-so. However, in response to this item, 21 students submitted the incorrect response do-re-mi-do-fa-so-fa. This accounted, in part, for the lower difficulty index obtained in item 9 as opposed to that obtained for item 10. The melody employed in item 10 was composed entirely of notes in step-wise movement (see Figure 2). As reported in Goolsby (1989), notes in a step-wise pattern are easier to read than notes which are spaced further apart. As such, it would appear that note reading is influenced by the pattern of the notes on the staff.

Items 11 and 12 (19 and 20)

Items 11 and 12 were designed to assess a student's ability to determine the placement of the tonal center, commonly designated as the home tone of a given melody.

Indices of Difficulty and Discrimination. The difficulty indices obtained for items 11 and 12 were .45 and .40, respectively. The point-biserial correlations were .40 and .57.



Note: Students were required to identify the notes in the first two measures of item 9 and the first measure of item 10.

Figure 2. Items 9 and 10: Melodic structure.

Items 13 and 14 (25 and 26)

Items 13 and 14 were designed to assess a student's ability to notate an aurally-presented phrase.

Indices of Difficulty and Discrimination. The difficulty index obtained for item 13, .12, was below the accepted guideline of .20. The difficulty index obtained for item 14, .21, approached this guideline. The point-biserial correlations were .29 for item 13 and .31 for item 14. These items were examined further in order to ascertain why they were so difficult.

Following Field Study 1, the melodies employed in Test Version 1 were replaced by motives written in 3/4 meter—thus reducing the number of unknown notes from 7 to 5. New melodies were based on familiar motives (i. a step-wise ascending pattern in a major tonality and ii. a minor tonic triad in an ascending and descending pattern). Based on the findings for Field Study 2, the notation of the rhythm written above the notes was clarified.

There was insufficient evidence obtained in Field Study 3 to determine why step-wise movement from 'do' to 'so' employed in item 13 was more difficult to notate than the minor-triadic formula la-do-mi-mi-do-la constructed for item 14. It was possible, however, that in the second item students attended to the mirror-like structure of the melody. On the other hand, more students may have responded correctly to the second question due to practise effects.

Items 21 and 22 (23 and 24)

Items 21 and 22 were designed to assess a student's ability to identify the sharps or flats in an aural presentation of a melody.

Indices of Difficulty and Discrimination. The difficulty indices obtained for items 21 and 22 were .48 and .49, respectively. The point-biserial index of discrimination for item 21 was .28. The point-biserial index of discrimination for item 22, .18, was below the accepted guideline of .20. It was noted that, in Field Study 1, the point-biserial index of discrimination for item 24 was .42. This observed variation in response may have been due to revisions to the instructions to students: in Field Study 1 students were instructed to choose the melody that was the same as the melody they hear; in Field Study 3 students were instructed to listen for the sharps and flats and choose the melody that is the same as they hear.

In item 21, students were required to identify a melody which had been altered by sharps. At $r_{pbis} = .28$, the correct response was functioning properly. Option A, which contained flats, had a discrimination index of $r_{pbis} = -.25$ and was chosen by 25% of the students. Similarly, Option B, with no sharps or flats, had a discrimination index of $r_{pbis} = -.07$ and was chosen by 27% of the students. It is not known why Option B did not discriminate between high-scoring and low-scoring students.

In item 22, students were required to identify a melody which had been altered by flats. At $r_{pbis} = .18$, the correct option was functioning at slightly below the suggested guidelines. Option B, which contained sharps, had a discrimination index of $r_{pbis} = -.08$ and was chosen by 32% of the students. Option C, with no altered tones, had a discrimination index of $r_{pbis} = -.14$ and was chosen by only 18% of the students. In item

22, it is possible that, due in part to practise effects, students have learned to stay away from the option with no altered tones in favor of the options with either sharps or flats.

Items 23 and 24 (21 and 22)

Items 23 and 24 were designed to assess a student's ability to identify a repeating melodic sequence.

Indices of Difficulty and Discrimination. The difficulty indices obtained for items 23 and 24 were .45 and .55, respectively. The point-biserial index of discrimination obtained for item 23, .11, was below the accepted guideline of .20. The index of discrimination obtained for item 24 was .39.

Both items 23 and 24 were revised prior to the formulation of Test Version 3. Data obtained in Field Study 1 revealed that, for item 23, both options A and C were correct. This oversight in melodic construction was accounted for in the revised item employed in Test Version 3. In addition, in Field Study 2, the melody composed for option C (item 23) was revised in order to avoid the sharp and the ledger line. For both items, revised directions were used to explain more specifically how the notated melodies were constructed.

Point-biserial correlations were examined in order to assess the internal functioning of the items. In item 23, the correct response (option C) discriminated positively. However, at $r_{pbis} = .11$, this coefficient was below the desired level. Option A, chosen by 20% of students, did not discriminate between the overall performance of those students who chose that option and those students who did not choose that option ($r_{pbis} = .00$). This seemed surprising considering that this option was dismissed easily by students in Field Study 2 because the melody constantly descended in the last two measures. Option B discriminated negatively ($r_{pbis} = -.09$). Contrary to expectations, this option, which differed from the correct response in terms of both contour and interval, was chosen by 35% of the students. In item 24, for option A, the correct response, the discrimination index was $r_{pbis} = .39$. Further, point-biserial correlations of -.15 and -.26, for options B (chosen by 13% of students) and C (chosen by 30% of students) respectively, indicated that these options were working properly.

Item Revision. It was not possible to discern why item 23 was more difficult than item 24. Considering the possibility that students did not understand how to answer items 23 and 24, the instructions were revised again. This change is illustrated for item 23 in Figure 3. Item 24 was revised similarly.

Test Version 3

Listen to each melody once. The first measure of each melody below begins with the same melodic **pattern** or **sequence.** Measures 2 and 3 are different. Circle the letter above the **one** melody which **best continues** the repeating melody.

Test Version 4

Look at the music shown below. The first measure of melodies **A, B,** and **C** all begin with the **same melodic pattern**. After that, each melody is different. Circle the letter above the **one** melody where the melodic pattern shown in the first measure is repeated in different places on the staff.

Figure 3. Item 23: Comparison of instructions to students in Test Versions 3 and 4.

APPENDIX H

TEST VERSION 4S

Music Reading Test
Version 4s- Protocol Analysis

This is a music reading test. All questions are about melody. You are requested to talk aloud as you answer these questions. Say out loud everything that comes into your head while you are answering the questions. I am interested in what you are doing while you figure out the answer, not just what the answer is. Try to explain everything that you used to solve the questions. Do not try to hurry- take your time. It is important that you explain everything that you are doing and thinking.

Before we begin the test, please answer the questions in the box below:

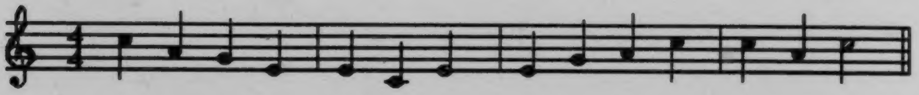
1. DO YOU TAKE PRIVATE MUSIC LESSONS? YES- _____ NO- ____

   IF ANSWER TO QUESTION 1 IS 'YES':

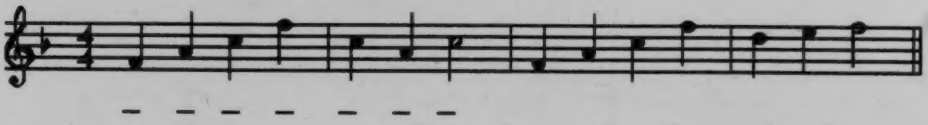
2. WHAT INSTRUMENT DO YOU PLAY? ________________________
3. HOW MANY YEARS HAVE YOU TAKEN PRIVATE LESSONS? ____

TURN THE PAGE TO QUESTION 1

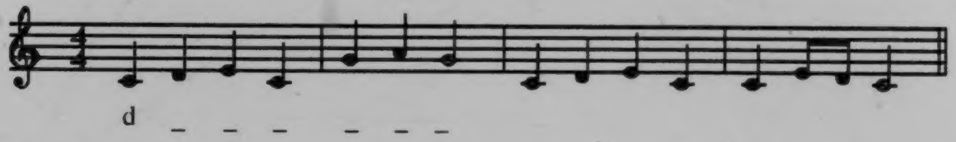
1. Look at the music shown below and circle all the pairs of notes that are separated by a **step**.



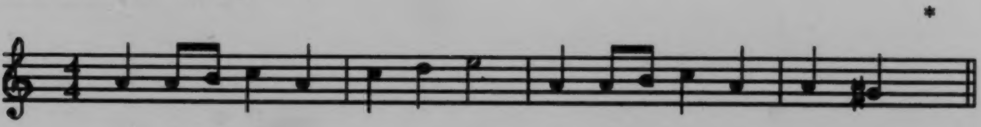
2. Look at the music shown below and write the **names of the notes** (A,B,C,D,E,F,G) on the lines under the notes.



3. Look at the music shown below and write the first letters of the **solfa syllables** (do,re,mi,fa,so,la,ti) on the lines under the notes. The first one is given.

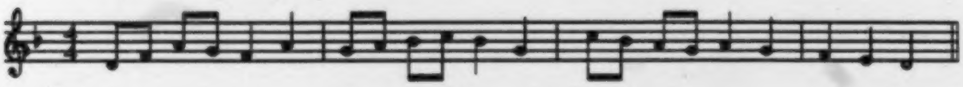


4. Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).

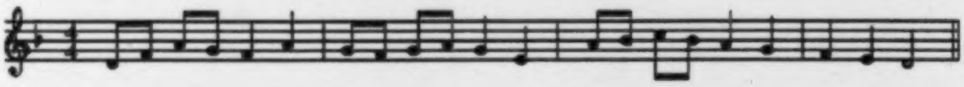


5. Listen to each melody once. The first measure of each melody below begins with the same melodic **pattern** or **sequence**. Measures 2 and 3 are different. Circle the letter above the **one** melody which **best continues** the repeating melody.

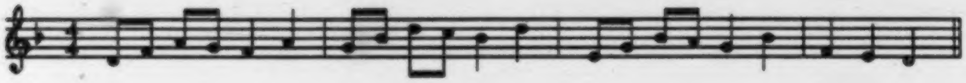
A.



B.



C.



APPENDIX I

FIELD STUDY 4: ITEM-BY-ITEM ANALYSIS

The 5 test items administered in Field Study 4 were analyzed on an item-by-item basis. For each item, the analysis of data is presented in three sections: first, an analysis of responses from students who answered the item correctly; second, an analysis of responses from students who answered the item incorrectly; and third, a comparison of data obtained in Field Study 2 and Field Study 4. Unless stated otherwise, item numbers refer to the order in which items were administered in Test Version 1.

Item 4.

Item 4 was designed to assess a student's ability to identify notes which were separated by a step.

Correct Response. Item 4 was answered correctly by students 2, 3, 6, 7, 8, and 9. All students solved this problem by referring to the spaces between the notes. Students 2, 3, 7, 8, and 9 described this separation in terms of steps. In the words of student 3, "They are separated by going up one note." Similarly, as described in the response of student 7, "One is on a line and then the line is going up to the space straight above it." Leaping notes, in this case the interval of a third, were identified by students 2, 3, and 9. Students 2, 3, 4, 5, and 9 referred to repeated notes. In the words of student 9, "The next two are on the same space." Students 3, 6, 7, 8, and 9 identified the spaces between the notes in terms of the letter names. "Those two letters [notes] are 'A' going down to 'G'," explained student 7.

Incorrect Response. Students 1, 4, 5, and 10 answered this item incorrectly. Student 1 based her response on the contour of the melody. Students 4 and 5 identified repeated notes in the apparent absence of attention to notes which formed a step or a leap. In both cases, it was possible that the ineffective problem-solving strategies employed by these students were due, in part, to an inability to understand the problem. Student 5 did not circle consecutive notes. Rather, she circled notes that repeated on the same line or space regardless of the intervening note (see Figure 1). Student 4 thought that steps were bar lines. The cue for this response may have come from the instructions which requested that students circle notes that were separated by a step. As revealed in the protocol, this student thought that the bar line was the separation referred to in these directions. As illustrated in Figure 2, this student's attention to repeated notes was due to their proximity.

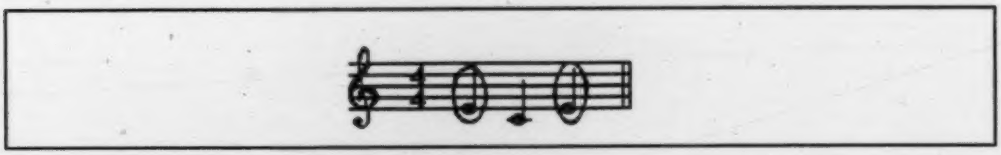


Figure 1. Stepping notes identified by student 5.

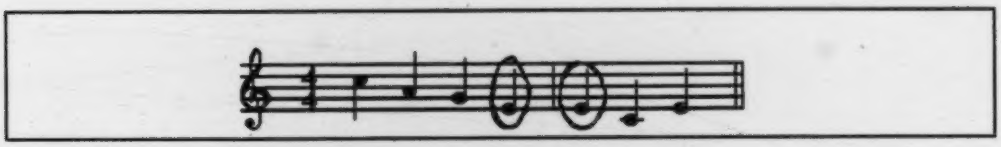


Figure 2. Stepping notes identified by student 4.

As noted above, it is possible that students 4 and 5 responded incorrectly because they did not refer to interval structure with relation to repeats, steps, and leaps. However, it must be noted that conceptualization in these terms did not guarantee a correct response. This is illustrated by student 10 who correctly described all interval types but mistakenly identified the interval of a third between middle 'C' and 'E' as a step. This confusion, attributed to middle 'C' on the ledger line, may indicate that intervals are more difficult to identify visually when a note is printed beyond the five-line staff. Further investigation is needed in order to justify this interpretation.

Comparison to Field Study 2. In contrast to the protocol data collected in Field Study 2, students were required to respond to this item in the absence of the aural stimulus. A comparison of the data obtained in Field Study 2 and Field Study 4 revealed that response strategies were similar regardless of the method of presentation. In both studies students described step-wise movement. Moreover, notes separated by a step were contrasted with repeating notes and leaping notes. In both investigations, students who described the overall contour of the melody, in the absence of attention to repeats, steps, or leaps, did not provide a correct response. It was noted that five students involved in Field Study 4 referred to note names while solving this problem. This strategy was not employed in Field Study 2. Correct identification of step-wise movement was not dependent upon use of this strategy. Therefore, due to the similarity of strategies employed in Field Study 2 and Field Study 4, this item was presented subsequently in visual form only (Field Study 5).

Item 7

Item 7 was designed to assess a student's ability to identify notes by their absolute letter names.

Correct Response. Students 2, 3, 5, 6, and 10 answered this item correctly. Students 3 and 5 used words or phrases to remember the names of the notes. Student 6 remembered the notes by using the note middle 'C' as a reference point. Student 10, who used the note 'E' on the bottom line as a reference point, responded incorrectly when he tried to "figure [it] out in [his] head." This student corrected his errors when requested to think aloud. Student 2 combined reference to the G clef with the order of the notes, "We are in [the] treble clef. It is a G clef, so the second line from the bottom is a 'G'. This one is seven steps so G-A-B-C-D-E-F." He extended his response further by referring to the 'F' major arpeggio as presented in the first measure of the melody. This extended response was due, most likely, to his background as a cellist.

Incorrect Response. Students 1, 4, 7, 8, and 9 answered this item incorrectly. Students 8 and 9 remembered the names of the notes by using the note middle 'C' as a reference point. In the words of student 8, "I'm counting up from middle 'C' up to the note—E-F-G-A." Students 8 and 9 combined this method with reference to the order of the notes, reference to repeated notes, and reference to lines and spaces (student 9 only). Apparently, errors in the response were due to an inability to equate the number of lines or spaces between notes when stated in alphabetical order with the number of lines or spaces between notes as presented visually on the staff. For example, student 8 correctly indicated that 'E' was three notes above 'C' but the interval indicated in the notation was a fourth—'C' to 'F'. Student 9 misrepresented the placement of notes on the staff, "On the very bottom there is a 'C' usually. It has a line through it. On the top you have a 'D'." Student 7 correctly named note 'B' on the middle line but was unable to relate the remaining notes to this reference point.

Students 1 and 4 referred to the order of the notes, recognized repeated notes, and attended to the relative highness of notes. For the most part, student 1 relied on her recorder experience, "I'm going to think back to when I play the recorder. The lowest one [is] 'G' because the 'G' makes the low sound." This strategy was unsuccessful because, from low to high, the notes as played on the recorder do not correspond to the notes as printed on the staff when 'G' is used as a reference point. Similarly, student 4 relied on general references to high and low notes, "That would be be a 'C' because I don't think 'F's are that high." As illustrated in these responses, the problem was not solved correctly by attending to the relative highness and lowness of notes in the absence of a more specific reference point.

Comparison to Field Study 2. In contrast to the protocol data collected in Field Study 2, students were required to respond to this item in the absence of the aural stimulus. A comparison of the data obtained in Field Study 2 and Field Study 4 revealed that response strategies were similar regardless of the method of presentation. In both studies, strategies for successful solutions included attention to the general order of the notes in terms of letter names, recognition of the intervals between notes, and observance of the lines and spaces on the staff. Although students in Field Study 4 relied less on memory devices (ie- words or phrases which correspond to the names of the notes) than students in Field Study 2, it appeared that effective solution of this problem was not dependent upon the aural perception of the melody. Therefore, in Field Study 5, this item was presented in visual form only.

Item 5

Item 5 was designed to assess a student's ability to equate notes with their corresponding solfa syllables.

Correct Response. Students 2, 3, 4, 5, 6, 8, 9, and 10 answered this item correctly. With the exception of student 8, these students related the order of the solfa syllables to the order of the notes on the staff. "It goes up notes. That's the bottom and that's the top and these spaces are going up like 'do' is the bottom and 'ti' is the top," explained student 10. With the exception of respondents 2 and 3, students referred to previous notes as they proceeded with the question. In the words of student 10, "Then you come down and you know that one's 'do' [because] its the same note [as] that one." Student 3 spoke of the Curwen handsigns associated with the syllables, a skill she learned in choir. Student 1 referred to the song *Do Re Mi* from the movie *The Sound of Music.*

Students 2, 4, 6, 9, and 10 indicated the spaces between notes in terms of intervals. "Since its a fifth ['C' to 'G'] we go from 'do' up to five—1, 2, 3, 4, 5—so it's a 'so'," explained student 2. Similarly, students 4, 5, and 10 referred to the lines and spaces of the staff. In the words of student 5, "The next one is on the line and its the next highest so it is 'mi'." Use of this strategy, however, contributed to the incorrect response submitted by student 1 who represented the space between 'E' and 'G' as the interval 'mi' to 'fa' rather than 'mi' to 'so'. Students 3, 6, 8, and 9 related the solfa syllables to the letter names of the notes. "Do-re-mi is on a scale if you go up from 'C'. 'C' would be 'do' and then then next one would be 're'," responded student 9.

Incorrect Response. Students 1 and 7 answered this item incorrectly. Student 1 referred to the order of the notes to the order of the solfa syllables but identified the interval 'do' to 'so' as the interval 'do' to 'fa'. Student 7 did not employ any strategies in the solution to this problem. In this student's words, "I don't understand any of it."

Comparison to Field Study 2. A comparison between responses obtained in Field Studies 2 and 4 indicate that, in general, students used similar techniques to solve this task as presented in this item regardless of whether or not they heard the melody. Strategies included attention to the general order of the syllables, reference to the letter names of the notes, recognition of the intervals between the notes, and observance of the lines and spaces on the staff. This item was presented subsequently in visual form only.

Item 19

Item 19 was designed to assess a student's ability to determine the placement of the tonal center, commonly designated as the home tone, in a given melody.

Key of a minor

Correct Response. Students 1, 2, 3, 8, and 9 answered this item correctly. Students 1, 3, 8, and 9 relied upon the home tone rule. "The first note is an 'A' so that means the home tone is 'A' and you can bring it over to the end and put it on 'A' again," explained student 9. Student 1 did not equate the terms home tone and tonal center, "The home tone would be the first note I guess. Or, if you wanted the tonal center, I would count the notes and get right [to] the middle note." She applied the home tone rule because that solution would be easier than looking for the center note. It must be recognized that the home tone, as defined by this rule, is reflected in the structure of the melody which, written in the 'a' minor diatonic scale, begins and ends on the note 'A'. Even though melodies may begin on any degree of the diatonic scale, it appeared that many students define the tonal center in terms of the first note of the melody. Student 2 related the notes of the melody to solfa syllables, "It would end on the 'A' because from 'ti' to 'do' would be ending the melody.

Incorrect Response. Students 4, 5, 6, 7, and 10 answered this item incorrectly. Of these, students 4, 5, 6, and 10 selected the final note by dividing the melody into two motives (see Figure 3). In the words of student 5, "All I'm going to do is look at the pattern. [It] keeps on going until it comes to the beginning again which is an 'A'." Students who used this strategy exclusively did not solve the task correctly. Moreover, interpretations based on this technique did not lead to consensus. Student 6 ended the melody on the note 'E,' "The second bar goes up and then I thought the fourth bar should go down." Student 5 thought the melody should rise to 'C' because there were two 'C's in the first motive but only one 'C' in the second motive. Student 10 analyzed the intervals between the notes and concluded that the last note should rise two spaces because the previous notes had fallen by one. Student 4 claimed that the melody should fall to 'E' but did not justify this decision. Student 7, the only respondent who did not apply the home tone rule or attend to the contour of the melody, tried to play the melody over in his head (audiation). He chose the note 'C' by "making a guess at it."

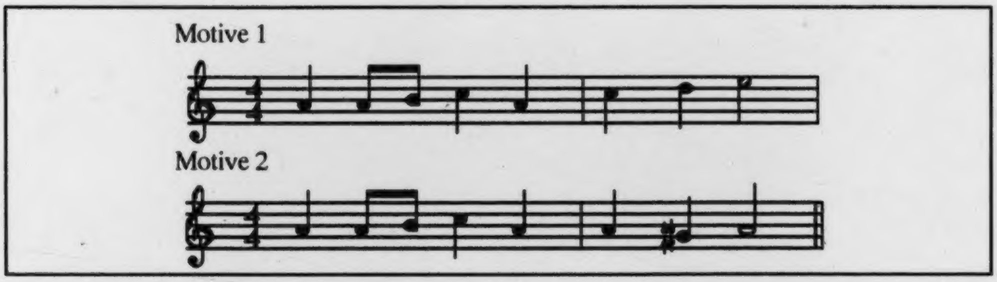
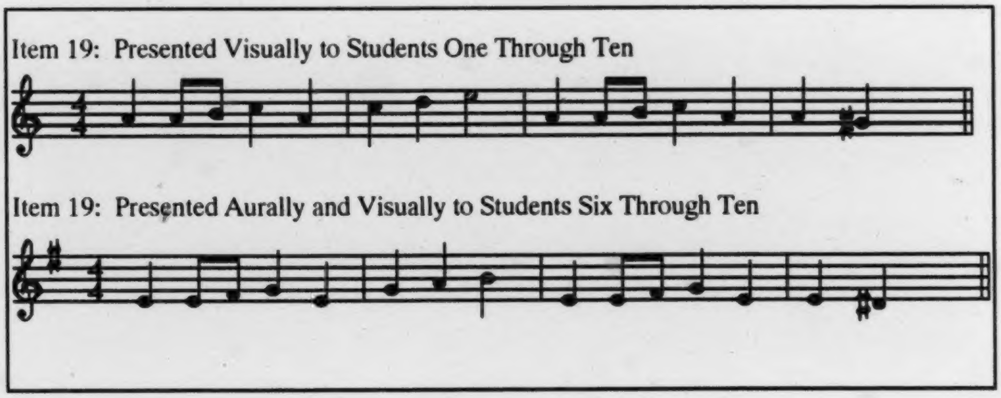


Figure 3. Item 19: Melodic structure.

Implications. During the first administrative session (students 1 through 5) it became possible that item 19 should be presented both aurally and visually. In order to examine this further, student 2 (who responded incorrectly by attending to the overall structure) and student 3 (who supplied the correct response by applying the home tone rule) heard the melody after they had completed the item. Upon hearing the melody, student 2 sang the correct note and changed his answer, "I wouldn't end it on 'F'. It would end on the 'A'." Student 3 hummed the correct ending and justified her answer, "That would sound good you know." As illustrated in Figure 4, this aspect of item development was explored further in the second administrative session (students 6 through 10) by means of a sixth item in which the melody employed in item 19, transposed to the 'e' minor diatonic scale, was presented to all students both aurally and visually.


Item 19: Presented Visually to Students One Through Ten

Item 19: Presented Aurally and Visually to Students Six Through Ten

Figure 4. Items 4 and 6: Comparison of melodic structure.

Key of e minor. Item 19 was administered to students 6, 7, 8, 9, and 10 in the key of 'e' minor.

Correct Response. Students 6, 8, 9, and 10 answered this item correctly. Student 6, 8, and 10 attended to the aural stimulus. In the words of student 8, "[I chose 'E'] because its what I think would sound the best." Student 10 extended his response further with reference to the E-D#-E motive, an ending common to Western tonal music, "It comes up to the same note as here to repeat it probably." Student 9 combined reliance on the home tone rule with reference to the aural stimulus, "It just sounds better if you would go up than going down." Considering misconceptions surrounding the home tone rule, attention to the aural stimulus is probably a more reliable strategy.

Incorrect Response. Student 7 answered this item incorrectly. He observed both the overall structure of the melody and the aural stimulus but was unable to correctly identify the tonal center.

Comparison of Item 19 when presented in the keys of 'a' minor and 'e' minor. All 5 students who heard this melody in the key of 'e' minor referred to the aural stimulus. With respect to students 8 and 9, attention to the aural stimulus for the item (a strategy employed when the melody was presented both aurally and visually) replaced, at least in part, reliance on the home tone rule (a strategy employed when the item was presented in visual form only). Students 6 and 10 answered the item incorrectly when presented with the 'a' minor melody in visual form only, but answered the item correctly when presented with the 'e' minor melody in both aural and visual form. Considering this, in Test Version 5, this item was presented both aurally and visually in the key of 'a' minor.

Item 11. Item 21 was designed to assess a student's ability to identify a repeating melodic sequence.

Correct Response. Students 2, 3, 4, 6, 7, 8, 9 and 10 answered this item correctly. Students 4 and 8 observed that the first measure of each option was the same. Students 2, 3, 4, 8, 9, and 10 compared the motives across options in terms of contour. "For [option] A and [option] C the notes go up and then they go down and then they go up and then go up and then they go on down and then up," explained student 3. Students 2, 3, and 8 compared the contour of the motives in terms of interval size. In the words of student 2, "Let's look at letter [option] A. The first three notes go up a third then fall down a second, then up a third—so that's our structure." Student 3 described these comparisons through reference to the names of the notes, "When you get to the third bar, [option] C goes from 'E' to 'G' and [option] A goes from 'A' to 'G'—so they're sort of going the opposite here." Student 9 attempted to complete the task by locating all the motives that began with the note 'D', the initial note of the first motive, "It starts with a 'D' and then you could go for the rest and see if you can find another bar where the beginning note is a 'D'." The student compared motives by contour after realizing that the first strategy would not lead to the correct response.

Student 6 described the motives in terms of lines and spaces. In the words of student 6, "[It's] number [option] C. [In] the first and the second and the third bar the first three notes start on lines and the fourth note starts on a space and the last two notes start on the lines."

Student 4 adapted strategies employed by the other students. He described how motives with different patterns could be rewritten to mirror the contour of the motive initially presented in each option, "These are mostly the same. All you've got to do is put it down one." In addition, he used his finger to measure the distance between notes realizing that if the motives were the same, then this distance would be the same.

Incorrect Response. Students 1 and 5 answered this item incorrectly. Student 1 described the rhythmic structure in terms of the direction of the note-tails. When asked why option A was incorrect, she replied, "They have a 'titi', another 'titi', and then two 'ta's' but the second 'titi' is upside down and then the first 'ta' is upside down." Student 5 observed that the first measure of each option was the same but, beyond this observation, did not know how to proceed with the question.

Comparison to Field Study 2. In contrast with the protocol data collected in Field Study 2, in Field Study 4 students were required to respond to this item in the absence of an aural stimulus. Responses obtained in Field Study 2 relied, for the most part, on contour and range (an aspect which was negated due to item revision). Similarly, in Field Study 4, strategies for the successful completion of the tasks included attention to contour, recognition of the intervals between notes, and observance of the lines and spaces on the staff. In consideration of this, it appeared that solutions were not dependent upon the aural perception of the melody. Therefore, in Test Version 6, this item was presented in visual form only.

APPENDIX J

TEST VERSION 5

Music Reading Test- 5
March, 1995

This is a music reading test. All questions are about melody. The test is divided into two parts- Part 1 and Part 2. The directions and music for Part 1 are on the cassette tape. Part 2 is completed silently.

Before the test begins, we will complete the questions in the box below:

1. DO YOU TAKE PRIVATE MUSIC LESSONS? YES-_____ NO-____

   IF ANSWER TO QUESTION 1 IS 'YES':

2. WHAT INSTRUMENT DO YOU PLAY? ________________________
3. HOW MANY YEARS HAVE YOU TAKEN PRIVATE LESSONS? ____

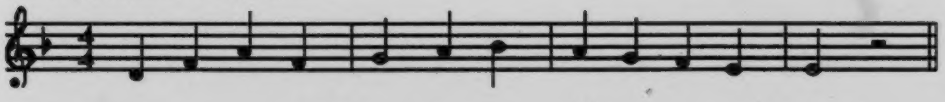
PART 1

The directions and melodies for **PART 1** are presented on this audio tape. Once the test has started, you must stay with the tape at all times. If you do not have time to complete a question, answer quickly and go on to the next one. If your pencil breaks, put up your hand and you will be given another one.

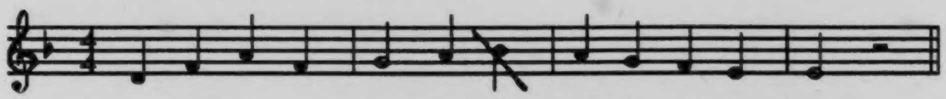
PART 1A- There are 3 practice questions and 10 test items in part 1a. For each question you write your answer on the printed music.

PRACTICE QUESTION 1

Listen to the melody once. Follow the music and put a line through the **last note** of the melody you hear sung aloud.

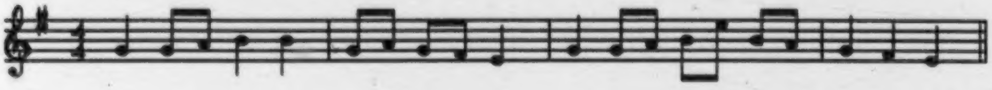


The answer to this question is shown below:

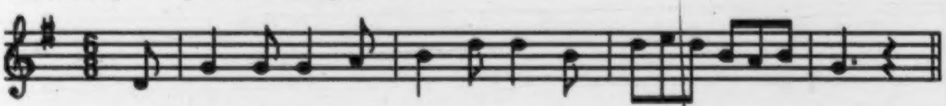


We will now stop the tape. Ask any questions you have about the directions.

1. Listen to the melody once. Follow the music and put a line through the **last note** of the melody that you hear **sung aloud**.

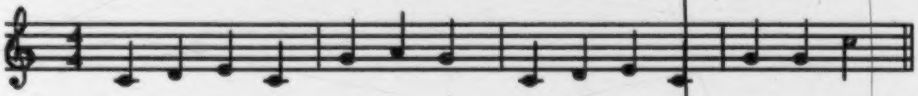


2. Listen to the melody once. Follow the music and put a line through the **last note** of the melody that you hear **sung aloud**.

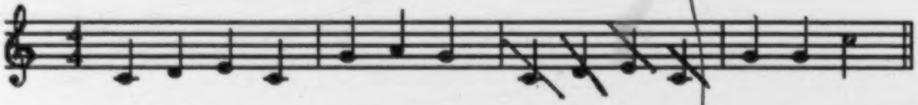


PRACTICE QUESTION 2

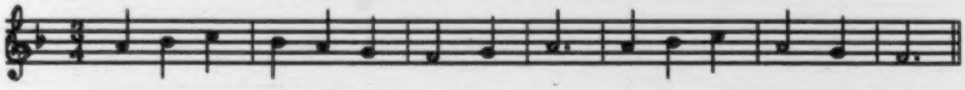
Only part of the melody written below is sung aloud once. As you listen to the melody, follow the music and put a line through the notes that are **not** sung aloud.



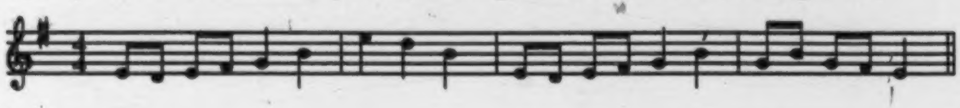
The answer to this question is shown below:



3. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the music and put a line through the notes that are **not** sung aloud.



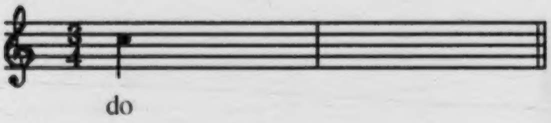
4. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the music and put a line through the notes that are **not** sung aloud.



PRACTICE QUESTION 3

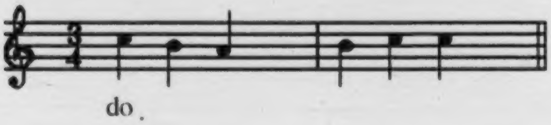
Listen to the melody three times. Write the notes of the melody on the staff below. The rhythm is shown above the melody in quarter notes. The first note of the melody is given.

rhythm: ♩ ♩ ♩ ♩ ♩ ♩



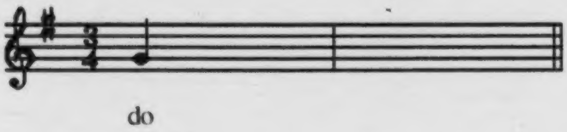
The answer to this question is shown below:

rhythm: ♩ ♩ ♩ ♩ ♩ ♩



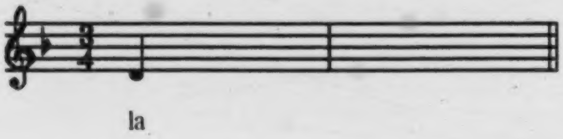
5. Listen to the melody three times. Write the notes of the melody on the staff below. The rhythm is shown above the melody in quarter notes. The first note of the melody is given.

rhythm: ♩ ♩ ♩ ♩ ♩ ♩

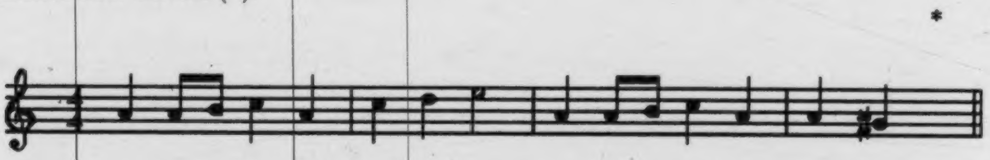


6. Listen to the melody three times. Write the notes of the melody on the staff below. The rhythm is shown above the melody in quarter notes. The first note of the melody is given.

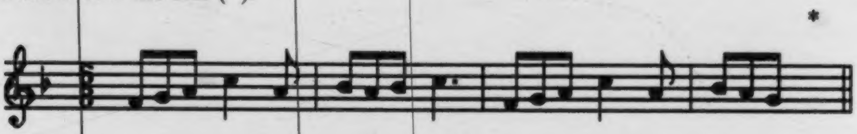
rhythm: ♩ ♩ ♩ ♩ ♩ ♩



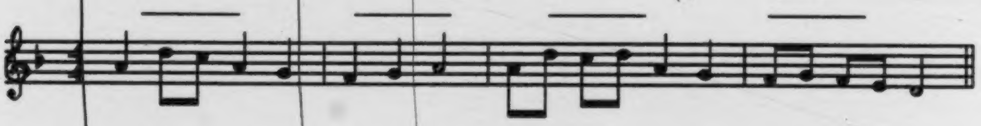
7. Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).



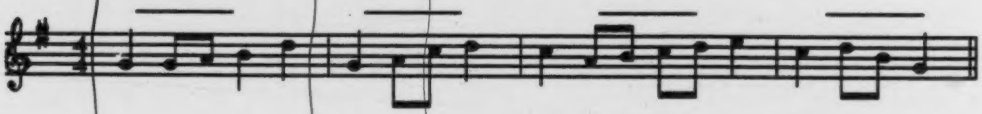
8. Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).



9. Listen to the melody twice. Place a check mark (√) on the line above the **one** measure which is **different** than the melody you hear.



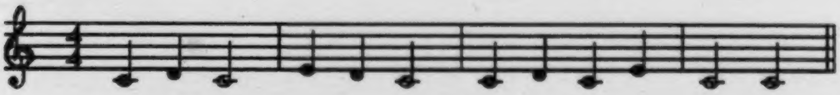
10. Listen to the melody twice. Place a check mark (√) on the line above the **one** measure which is **different** than the melody you hear.

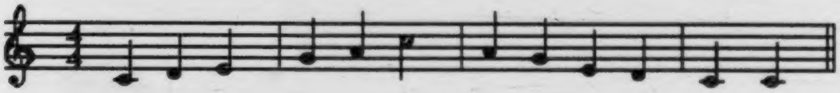


PART 1B- There are 8 test items in part 1B. Each question is multiple-choice. You decide on the correct answer from three choices- melody A, melody B, or melody C.

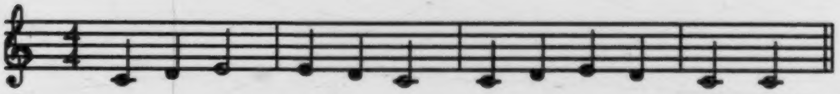
11. Listen to the melody twice. Circle the letter above the **one** melody which is the **same** as the melody you hear.

A



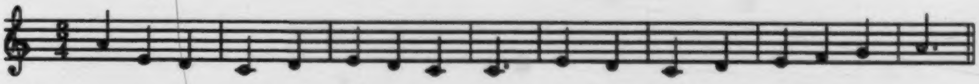


C

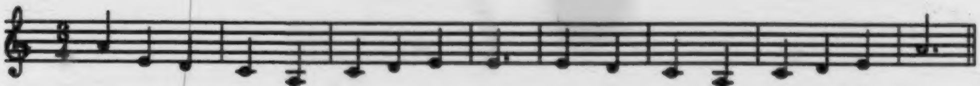


12. Listen to the melody twice. Circle the letter above the **one** melody which it the **same** as the melody you hear.

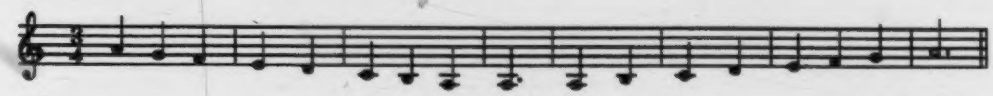
A



B

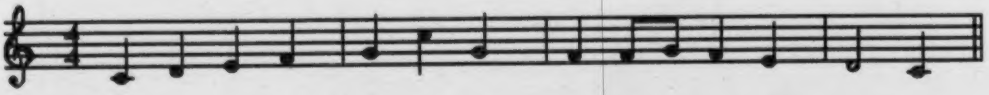


C

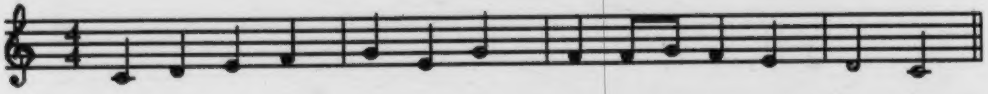


13. Listen to the melody twice. Circle the letter above the one melody which is the same as the melody you hear.

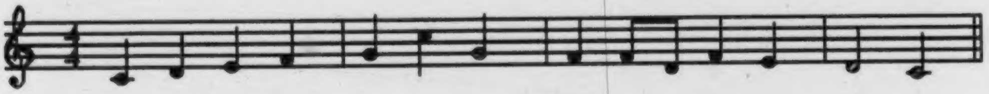
A.



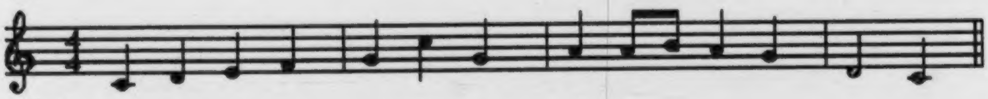
B.



C.

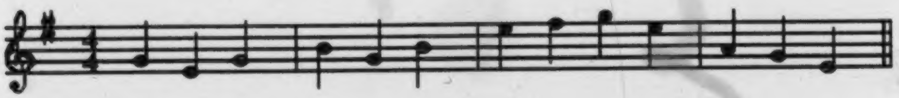


D.

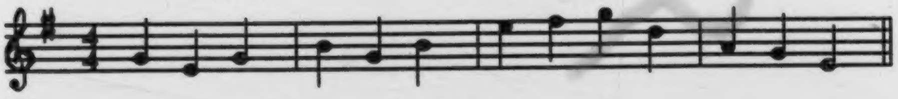


14. Listen to the melody twice. Circle the letter above the one melody which is the same as the melody you hear.

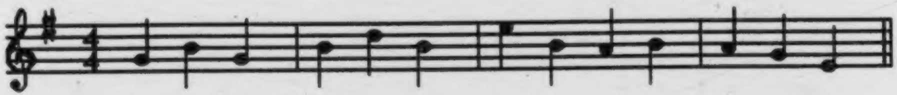
A.



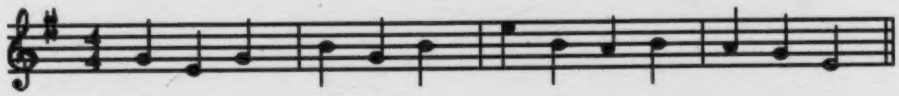
B.



C.

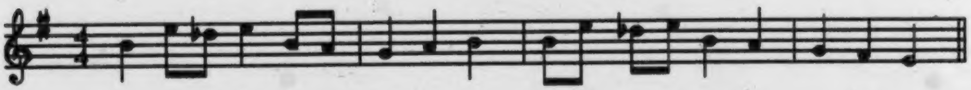


D.

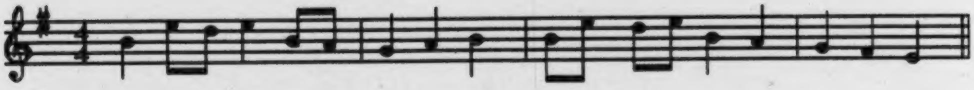


15. Listen to the melody twice. Listen for the sharps or flats and circle the letter above the **one** melody which is the **same** as the melody you hear.

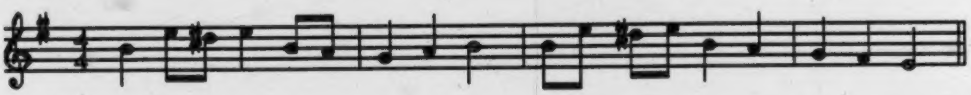
A.



B.

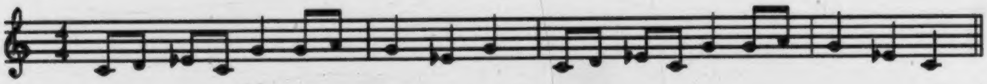


C.

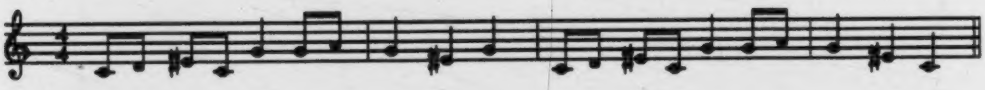


16. Listen to the melody twice. Listen for the sharps or flats and circle the letter above the **one** melody which is the **same** as the melody you hear.

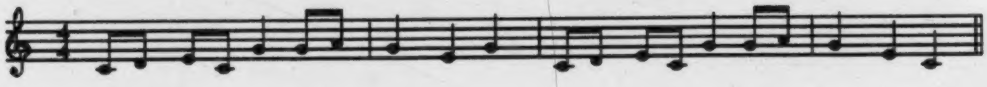
A.



B.



C.

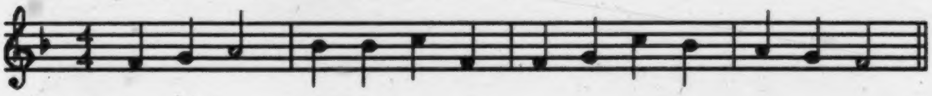


P

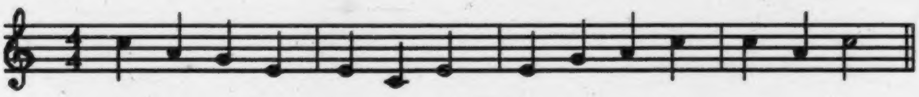
PART 2

There are 8 questions in part 2. They are all answered silently. You have 10 minutes to complete the test.

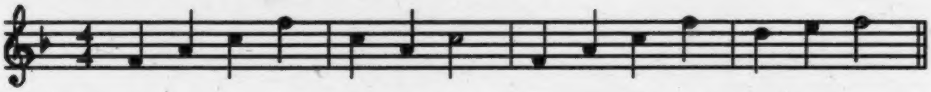
17. Look at the music shown below and circle all the pairs of notes that are separated by a **leap**. A leap is sometimes called a skip.



18. Look at the music shown below and circle all the pairs of notes that are separated by a **step**.

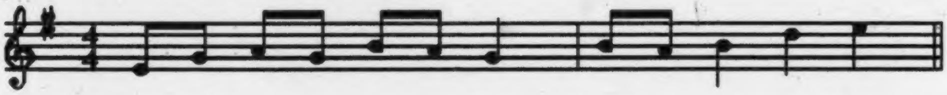


19. Look at the music shown below and write the **names of the notes** (A,B,C,D,E,F,G) on the lines under the notes.



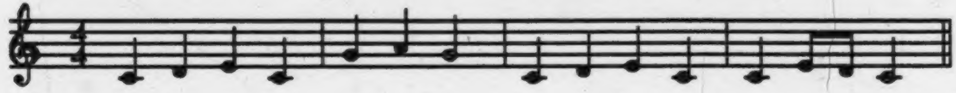
_ _ _ _ _ _ _

20. Look at the music shown below and write the **names of the notes** (A,B,C,D,E,F,G) on the lines under the notes.



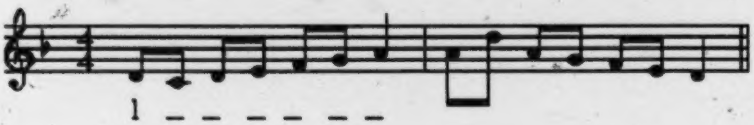
_ _ _ _ _ _ _

21. Look at the music shown below and write the first letters of the **solfa syllables** (do,re,mi,fa,so,la,ti) on the lines under the notes. The first one is given.



d _ _ _ _ _ _

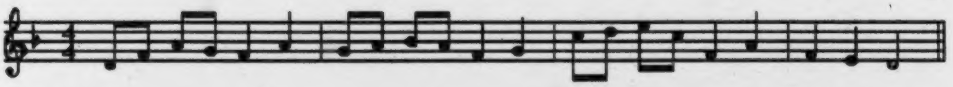
22. Look at the music shown below and write the first letters of the **solfa syllables** (la,ti,do,re,mi,fa,so) on the lines under the notes. The first one is given.



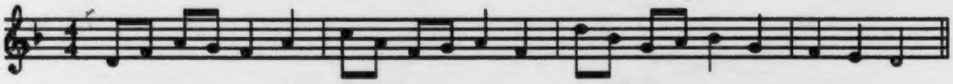
l _ _ _ _ _ _ _

23. Look at the music shown below. The first measure of melodies **A, B,** and **C** all begin with the **same melodic pattern**. After that, each melody is different. Circle the letter above the **one** melody where the melodic pattern shown in the first measure is repeated in different places on the staff.

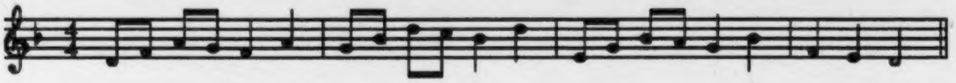
A.



B.



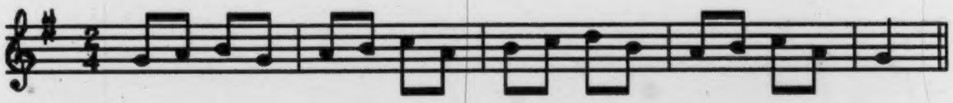
C.



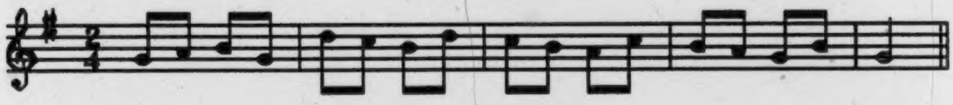
TURN THE PAGE TO QUESTION 24

24. Look at the music shown below. The first measure of melodies **A, B,** and **C** all begin with the **same melodic pattern**. After that, each melody is different. Circle the letter above the **one** melody where the melodic pattern shown in the first measure is repeated in different places on the staff.

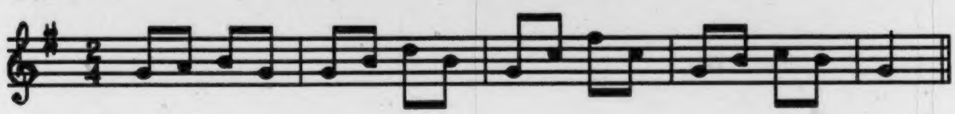
A.



B.



C.



APPENDIX K

FIELD STUDY 5: ITEM-BY-ITEM ANALYSIS FOR ITEMS NOT REVISED

Test Version 5 was administered to students in 2 grade 5 classes in March, 1995. All students attended one of two schools within School District 3 and each class was taught by one of two music specialists. Of the 24 items administered in Field Study 5, 11 items were not revised (items 3, 7, 8, 11, 12, 13, 14, 19, 20, 21, and 22). The item-by-item analysis for each cluster or pair of items is described below. The exception is test item 3, which is discussed in the absence of item 4 (an item which was revised). Unless specified otherwise, item numbers refer to the order in which items were administered in Test Version 5. Numbers in parentheses refer to the order in which the items were administered in Field Study 3.

Item 3 (Item 3 in Field Study 3)

Item 3 was designed to assess a student's ability to follow music notation. In these questions, the aural stimulus was divided into three sections—the first and last were sung aloud and the second was silent. Students were required to follow the notation and indicate which notes were silent.

Indices of Difficulty and Discrimination. For item 3, the difficulty index was .61. the point-biserial correlation was .52.

Items 7 and 8 (Items 11 and 12 in Field Study 3).

Items 7 and 8 were designed to assess a student's ability to determine the placement of the tonal center, commonly designated as the home tone of a given melody.

Indices of Difficulty and Discrimination. The difficulty indices for items 7 and 8 were .35 and .65. The point-biserial correlations were .29 and .28, respectively.

The difference in difficulty between the two items may be due to the structure of the melodies employed in the items. Students in this sample were better able to complete the melody in item 8, a major melody which descended to the tonic as opposed to item 7, a minor melody which ended with the A-G#-A (la-si-la) motive. It was noted that the apparent easiness of item 8 may be attributed to practise effects.

Items 11 and 12 (Items 17 and 18 in Field Study 3)

Items 11 and 12 were designed to assess a student's ability to identify the contour of a given melody.

Indices of Difficulty and Discrimination. The difficulty indices for items 11 and 12 were .87 and .65. The point-biserial correlations were .25 and .28.

As described in Field Study 3, high levels of difficulty were related to the nature of the task. In consideration of the data obtained in the former study, linear melodies were replaced by melodies with several changes in contour. This revision did not appear to change the pattern of response. In Field Study 5, the options which portrayed a complex contour discriminated between students who chose the option and students who did not choose the option having obtained point-biserial correlations of -.25 and -.31. These options were chosen by 9% and 22% of the students, respectively. In all probability, the increase in the percentage of students who chose the melody with the complex contour in item 12 was related to the nature of the task requirements—to choose the M-shaped melody.

In item 11, students were required to identify a V-shaped melody. The M-shaped melody, an incorrect option, had a discrimination index of $r_{pbis} = -.20$. As in the previous group administration, this item was chosen by only 4% of the students and, as such, was a non-functioning option. This, in combination with the rarity with which students chose the option representing the complex contour, contributed to the high difficulty index obtained for this item.

In item 12, students were required to identify an M-shaped melody. In this case the V-shaped melody, an incorrect option, did not discriminate between high-scoring and low-scoring students ($r_{pbis} = .02$). On the other hand, a point-biserial index of correlation of -.31 was obtained for option A (complex contour), indicating that low-scoring students were attracted to this option.

Items 13 and 14 (Items 19 and 20 in Field Study 3)

Items 13 and 14 were designed to assess a student's ability to identify the notational equivalent of an aural melody.

Indices of Difficulty and Discrimination. The difficulty indices for items 13 and 14 were .61 and .57. Differences between indices of difficulty obtained in Field Study 5 and Field Study 3 (item 19: $p = .83$; item 20: $p = .42$) were attributed to item revisions. More specifically, the items administered in Field Studies 3 and 5 contained 3 and 4 options, respectively. Additionally, for item 13 (19), the transformation in contour employed for the melody in option C was revised in order to avoid a note on a ledger line. In terms of interval, the differences between corresponding notes in option A (the correct response) and option B (an incorrect alternative) were reduced to intervals of a third. For item 14 (20), the difference between option C (the correct response) and option A (an incorrect alternative) was increased from one notes to two notes. In addition, the difference between the aural stimulus and the notational representation was accentuated by a high 'E' in the music notation.

In item 13, the point-biserial correlation for the correct response (option A), .47, exceeded the value of .37 obtained for item 19 in Field Study 3. It seems surprising that option B, the option in which the contour differed from the correct response in a different measure than that in which the other two incorrect alternatives differed from the correct response, would discriminate positively ($r_{pbis} = .24$). This was not considered problematic because the option was chosen by only 4% of the students. From a psychological perspective it seemed surprising that option C, which differed from the correct response in terms of contour, was chosen by 30% of the students. However, at $r_{pbis} = -.65$, it makes sense that this option discriminated between students to such a high degree. It seemed to follow that option D (which differed from the correct response in terms of interval) would discriminate positively ($r_{pbis} = .12$). It was surprising that this option was chosen by only 4% of the students.

In item 14 the point-biserial correlation for the correct response (option D), -.06, was lower than the value of .26 obtained for item 20 in Field Study 3. Moreover, the point-biserial correlation obtained for item 14 was below the accepted guidelines of .20. In consideration of this, the internal structure of the item was examined further to ascertain why this item discriminated negatively. As expected, option B (r_{pbis} = -.07), the option in which the contour differed from the correct response in the same measure as that in which option A differed from the correct response in terms of interval, was chosen by only 4% of the students. It followed that option A which discriminated negatively (r_{pbis} = -.15) and was chosen by 25% of the students, was a more popular option than option B. It is surprising that option C, the option in which the contour differed from the correct response in measures 1 and 2 (options A and B differed from the correct response in measure 3), discriminated positively (r_{pbis} = .33) and was chosen by 13% of the students.

Items 19 and 20 (Items 7 and 8 in Field Study 3)

Items 19 and 20 were designed to assess a student's ability to identify notes by their absolute letter names (A, B, C, D, E, F, and G).

Indices of Difficulty and Discrimination. The difficulty levels for items 19 and 20 were .61 and .83. Point-biserial correlations for both items were .47.

Items 21 and 22 (Items 9 and 10 in Field Study 3)

Items 21 and 22 were designed to assess a student's ability to equate notes with their corresponding solfa syllable names (do, re, mi, fa, so, la, and ti).

Indices of Difficulty and Discrimination. The difficulty indices for items 21 and 22 were .57 and .39. The point-biserial correlations were .33 and .28.

APPENDIX L

TEST VERSION 6

Music Reading Test- 6
April, 1995

This is a music reading test. All questions are about melody. The test is divided into two parts- Part 1 and Part 2. The directions and music for Part 1 are on the cassette tape. Part 2 is completed silently.

Before the test begins, we will complete the questions in the box below:

1. DO YOU TAKE PRIVATE MUSIC LESSONS? YES- _____ NO- ____

   IF ANSWER TO QUESTION 1 IS 'YES':
2. WHAT INSTRUMENT DO YOU PLAY? ________________________
3. HOW MANY YEARS HAVE YOU TAKEN PRIVATE LESSONS? ____

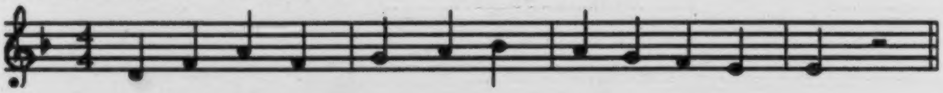
PART 1

The directions and melodies for **PART 1** are presented on this audio tape. Once the test has started, you must stay with the tape at all times. If you do not have time to complete a question, answer quickly and go on to the next one. If your pencil breaks, put up your hand and you will be given another one.

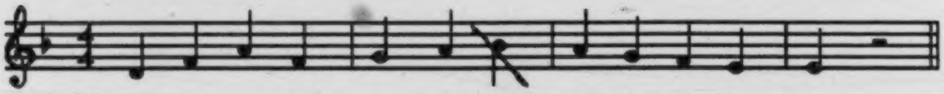
PART 1A- There are 2 practice questions and 8 test items in part 1a. For each question you write your answer on the printed music.

PRACTICE QUESTION 1

Listen to the melody once. Follow the music and put a line through the **last note** of the melody that you hear sung aloud.

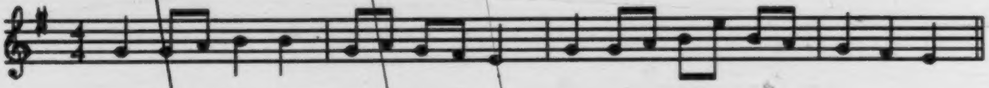


The answer to this question is shown below:

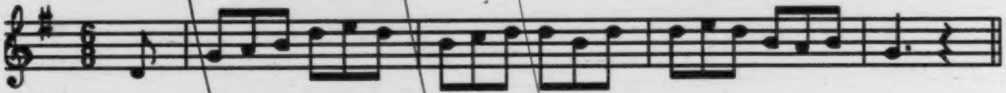


We will now stop the tape. Ask any questions you have about the directions.

1. Listen to the melody once. Follow the music and put a line through the **last note** of the melody that you hear **sung aloud**.

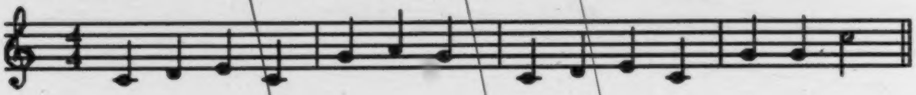


2. Listen to the melody once. Follow the music and put a line through the **last note** of the melody that you hear **sung aloud**.

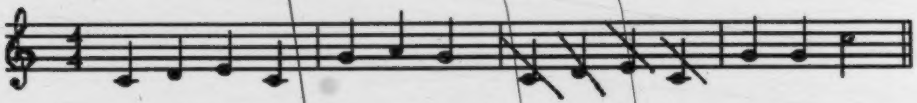


PRACTICE QUESTION 2

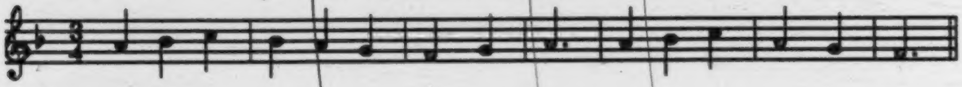
Only part of the melody written below is sung aloud once. As you listen to the melody, follow the music and put a line through the notes that are **not** sung aloud.



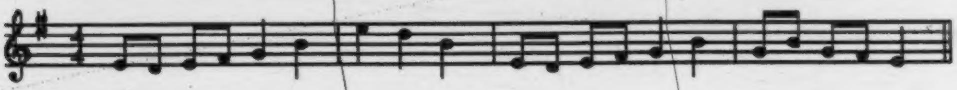
The answer to this question is shown below:



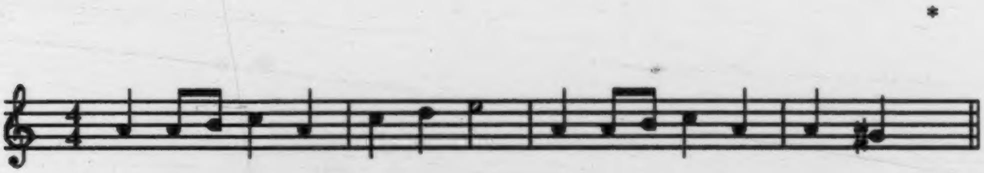
3. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the music and put a line through the notes that are **not** sung aloud.



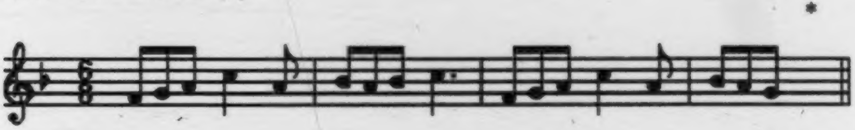
4. Only part of the melody written below is sung aloud once. As you listen to the melody, follow the music and put a line through the notes that are **not** sung aloud.



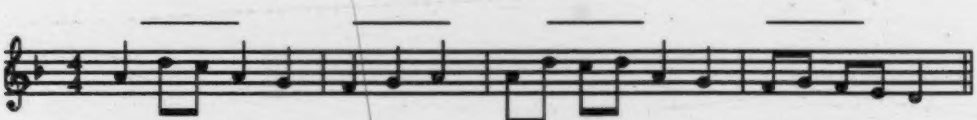
5. Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).



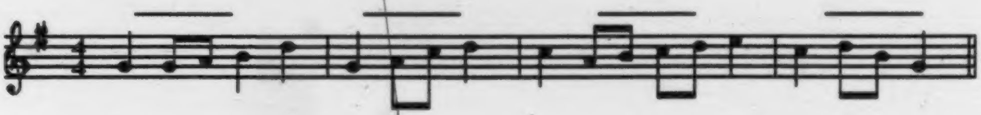
6. Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).



7. Listen to the melody twice. Place a check mark (√) on the line above the **one** measure which is **different** than the melody you hear.



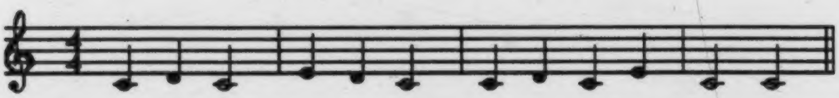
8. Listen to the melody twice. Place a check mark (√) on the line above the **one** measure which is **different** than the melody you hear.



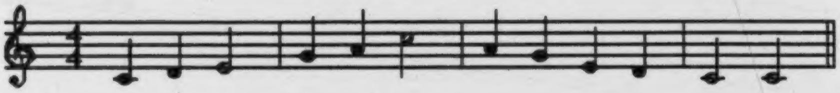
PART 1B- There are 6 test items in part 1B. Each question is multiple-choice. You decide on the correct answer from three choices- melody A, melody B, or melody C.

9. Listen to the melody once. Circle the letter above the one melody which is the **same** as the melody you hear.

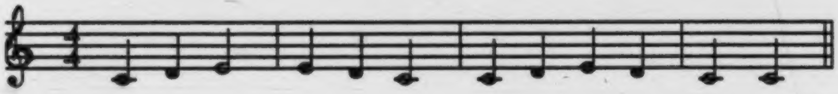
A



B.

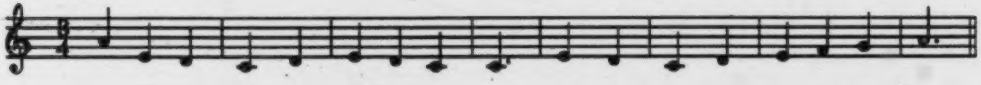


C

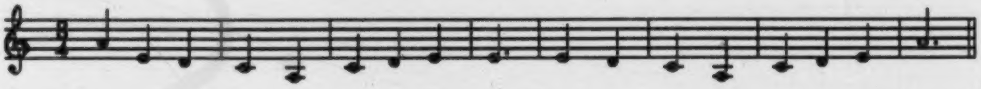


10. Listen to the melody once. Circle the letter above the one melody which it the **same** as the melody you hear.

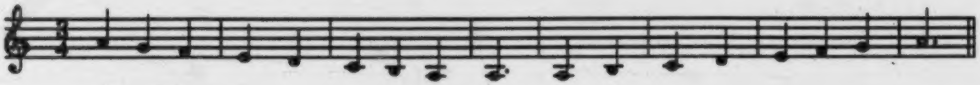
A



B

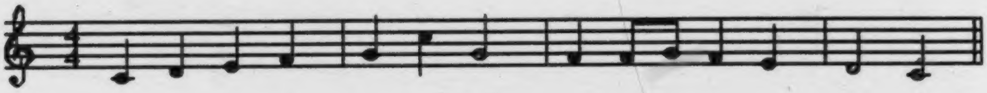


C

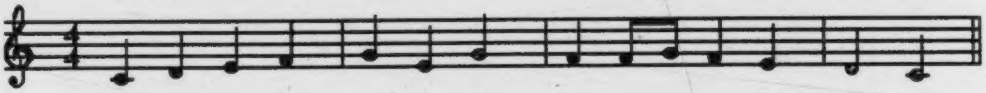


11. Listen to the melody twice. Circle the letter above the one melody which is the **same** as the melody you hear.

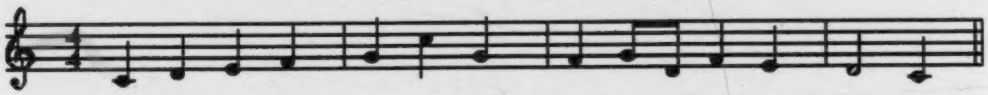
A.



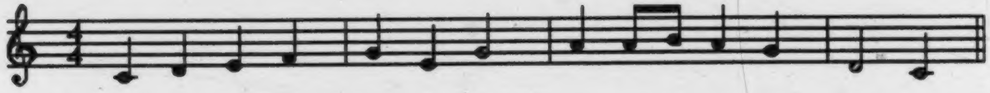
B.



C.

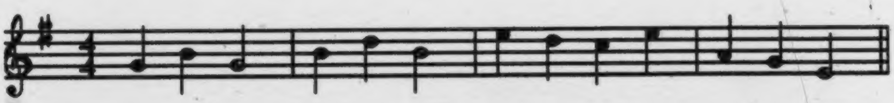


D.

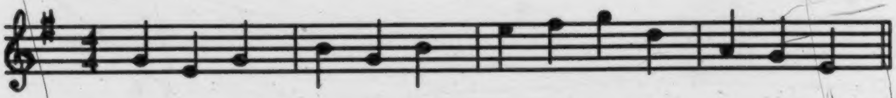


12. Listen to the melody twice. Circle the letter above the one melody which is the **same** as the melody you hear.

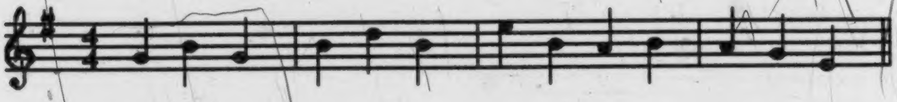
A.



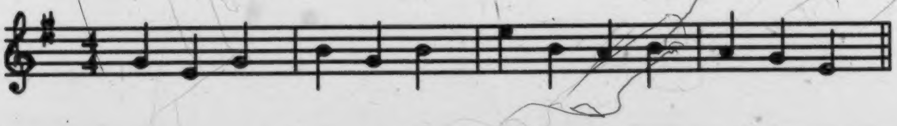
B.



C.

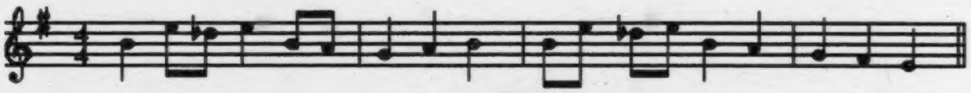


D.

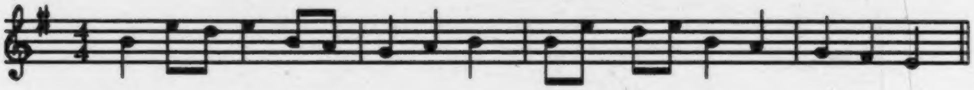


13. Listen to the melody twice. Circle the letter above the one melody which is the **same** as the melody you hear.

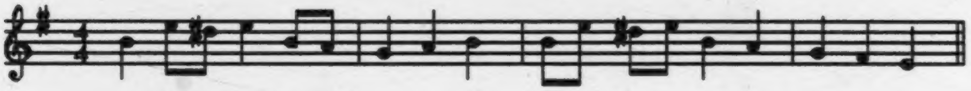
A.



B.

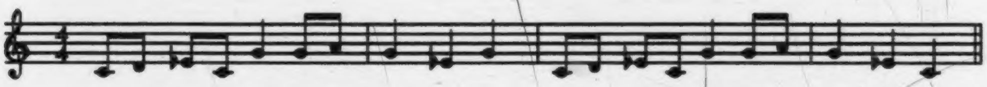


C.

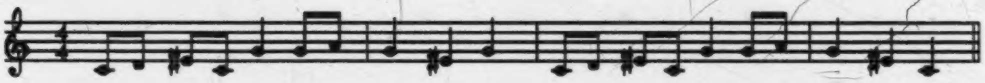


14. Listen to the melody twice. Circle the letter above the one melody which is the **same** as the melody you hear.

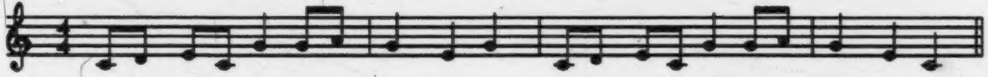
A.



B.



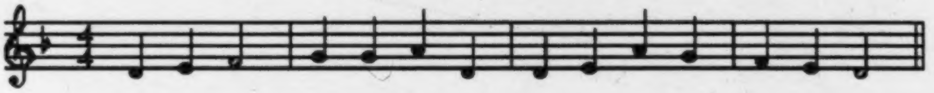
C.



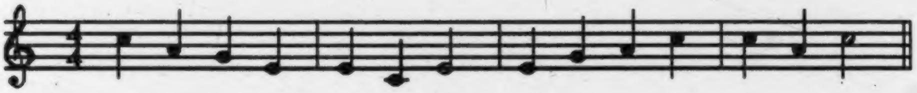
PART 2

There are 8 questions in part 2. They are all answered silently. You have 10 minutes to complete the test.

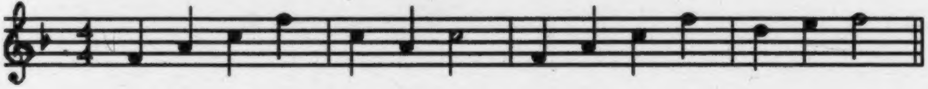
15. Look at the music shown below and circle all the places in the music where two notes beside each other are separated by a **leap**. A leap is sometimes called a skip.



16. Look at the music shown below and circle all the places in the music where two notes beside each other are separated by a **step**.

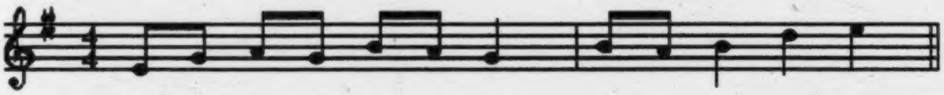


17. Look at the music shown below and write the **names of the notes** (A,B,C,D,E,F,G) on the lines under the notes.



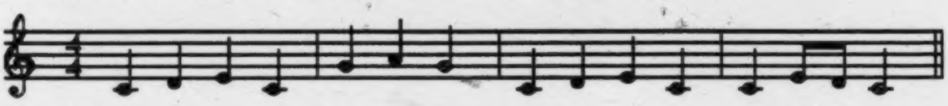
_ _ _ _ _ _ _

18. Look at the music shown below and write the **names of the notes** (A,B,C,D,E,F,G) on the lines under the notes.



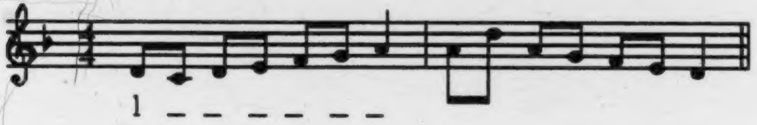
_ _ _ _ _ _ _

19. Look at the music shown below and write the first letters of the **solfa syllables** (do,re,mi,fa,so,la,ti) on the lines under the notes. The first one is given.



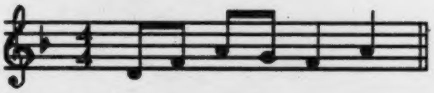
d _ _ _ _ _ _

20. Look at the music shown below and write the first letters of the **solfa syllables** (la,ti,do,re,mi,fa,so) on the lines under the notes. The first one is given.



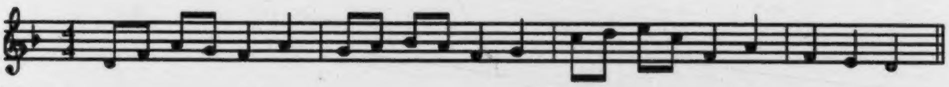
l _ _ _ _ _ _

21. Look at the following melodic pattern

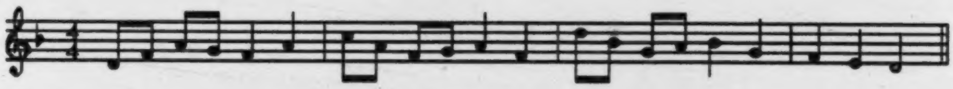


Circle the letter above the **one** melody where this melodic pattern is repeated in different places on the staff.

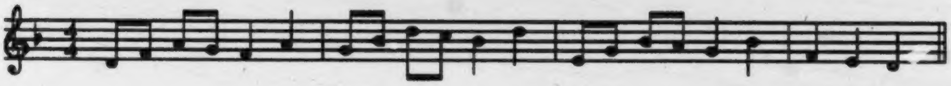
A.



B.

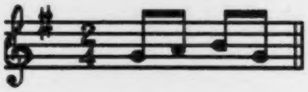


C.



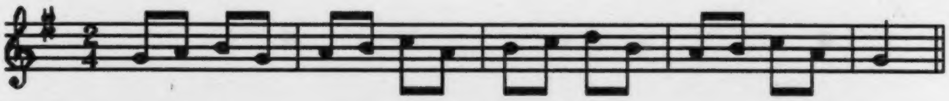
TURN THE PAGE TO QUESTION 22

22. Look at the following melodic pattern.

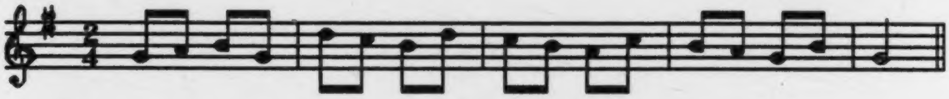


Circle the letter above the **one** melody where this melodic pattern is repeated in different places on the staff.

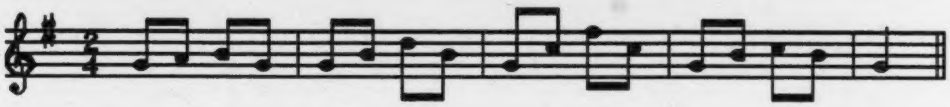
A.



B.



C.



APPENDIX M

FIELD STUDY 6: ITEM-BY-ITEM ANALYSIS FOR ITEMS NOT REVISED

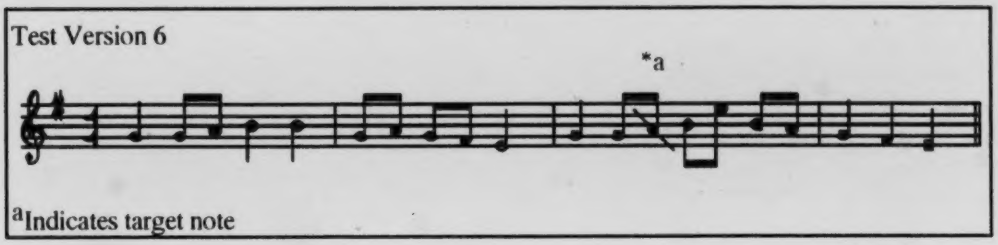
Test Version 6 was administered to 214 students in seven schools. All students attended a school within School District 1, 2, or 3 and each class was taught by one of seven music specialists. Of the 23 items administered in Field Study 5, 17 items were not revised (items 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 17, 18, 19, and 20). The item-by-item analysis for each cluster or pair of items is described below. The exceptions are items 12, 14, and 15 which are described in the absence of items 11, 13, and 16 (items for which the analysis is provided in Chapter 8). Item numbers refer to the order in which items were administered in Test Version 6.

Items 1 and 2

Items 1 and 2 were designed to assess a student's ability to follow music notation. In these questions, only part of the melody was sung aloud. Students were required to draw a line through the final note they heard.

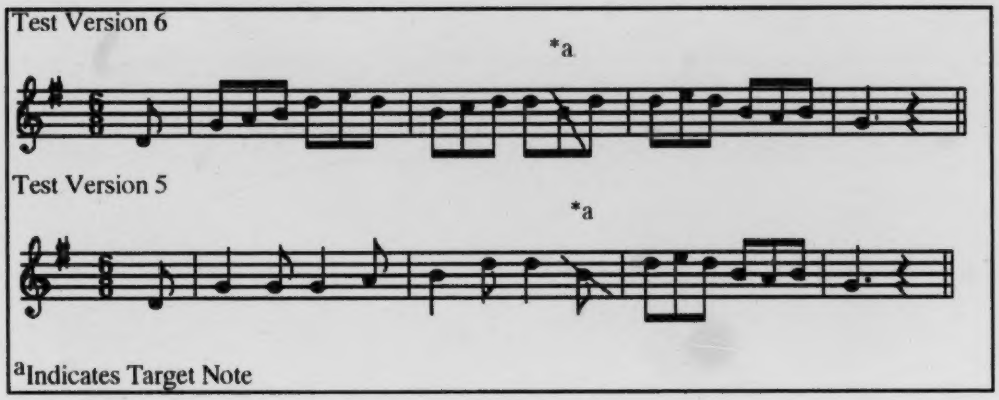
Indices of Difficulty and Discrimination. The difficulty indices of items 1 and 2 were .73 and .50, respectively. Both items discriminated positively (item 1: $r_{pbis} = .45$; item 2: $r_{pbis} = .39$). A comparison of difficulty indices obtained for items 1 and 2 in Field Study 5, .83, indicates that revisions intended to make the items more difficult had the intended effect. For item 1, in Field Study 5 students were required to identify a note on the first beat on measure three. In revised form, students were required to identify a note on a musically less important beat—the second half of the second beat in the third measure. In item 2, the quarter note-eighth note structure heard in Test Version 5 was replaced by a series of eighth notes.

In the two previous test versions items 1 and 2 were of equal difficulty. In consideration of this, it was necessary to examine the musical structure of these items in order to suggest why, as revealed in the data obtained in Field Study 6, item 1 was easier than item 2. As illustrated in Figure 1, it is possible that, in response to item 1, students used the half note on the second two beats of measure 2 and the quarter note on the first beat of measure 3 as reference points. Assuming this response strategy, item 2 was more difficult because the target note was the twelfth tone in a succession of 19 eighth notes (see Figure 2).


Test Version 6

[a]Indicates target note

Figure 1. Item 1: Target note.


Test Version 6

Test Version 5

[a]Indicates Target Note

Figure 2. Item 2: Comparison of target notes in Test Versions 5 and 6.

Items 3 and 4

Items 3 and 4 were designed to assess a student's ability to follow music notation. In these questions, the aural stimulus was divided into three sections—the first and last were sung aloud and the second was silent. Students were required to follow the notation and indicate which notes were silent.

Indices of Difficulty and Discrimination. The difficulty indices for items 3 and 4 were .64 and .36. Point-biserial indices of correlation, .58 and .36 obtained for items 3 and 4 respectively, were within suggested guidelines. The lower index of difficulty obtained for item 4 in Field Study 6 as compared to the index of difficulty obtained for this item in Field Study 5 (p = .74) was most likely due to the rhythmic structure of the silent section (see Figure 3). In Test Version 5 it appeared that the half note in measure 2 functioned as a reference point marking the final note in section 1. In revised form (Test Version 6) a shift in the placement of section 2 no longer allowed for this response strategy.

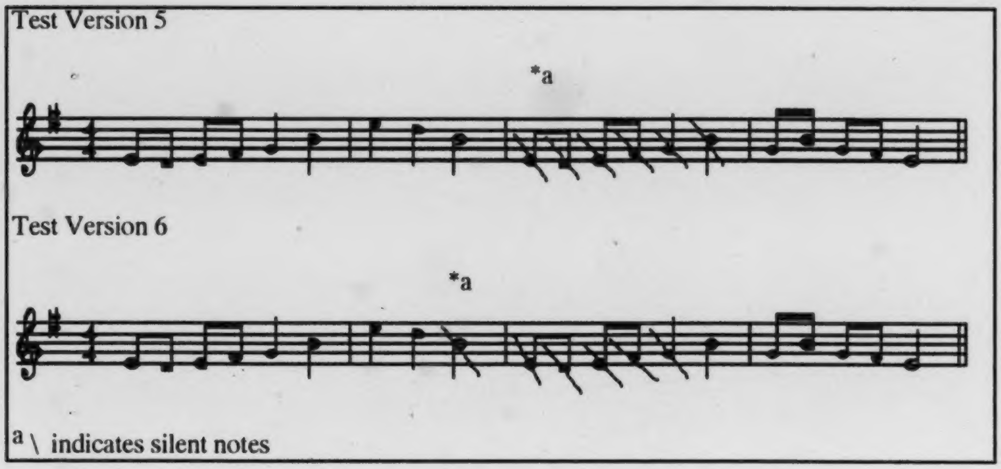


Figure 3. Item 4: Comparison of target notes in Test Versions 5 and 6.

Items 5 and 6

Items 5 and 6 were designed to assess a student's ability to determine the placement of the tonal center, commonly designated as the home tone of a given melody.

Indices of Difficulty and Discrimination. The difficulty index of items 5 and 6 were .43 and .62. Point-biserial indices of discrimination were .55 and .52.

Items 7 and 8

Items 7 and 8 were designed to assess a student's ability to identify the one measure of a given melody in which the aural presentation differed from the notational representation.

Indices of Difficulty and Discrimination. The difficulty indices for items 7 and 8 were .47 and .54. For items 7 and 8 respectively, point-biserial indices of correlation were .54 and .48 and all incorrect options discriminated negatively.

Items 9 and 10

Items 9 and 10 were designed to assess a student's ability to identify the overall contour of a given melody.

Indices of Difficulty and Discrimination. The difficulty indices for items 9 and 10 were .88 and .45. For items 9 and 10 respectively, point-biserial indices of correlation were .39 and .47 and all incorrect options discriminated negatively.

Item 12

Items 11 and 12 were designed to assess a student's ability to identify the notational equivalent of an aural melody when presented with three options in notational form.

Indices of Difficulty and Discrimination. The difficulty level for item 12 was .60. The point-biserial index of correlation was .56 and all incorrect options discriminated negatively.

Items 14

Item 14 was designed to assess a student's ability to identify the sharps or flats in an aural presentation of a melody.

Indices of Difficulty and Discrimination. The difficulty index for item 14 was .47. The point-biserial index of correlation for the correct response (option A) was .33. It is interesting to note that both incorrect options, chosen by 28% and 25% of the students, had indices of discrimination of -.19. In consideration of this, it appears that low scoring students were more likely to choose either of the two incorrect options. Thus, it appears that this item is functioning properly.

Item 15

Item 15 was designed to assess a student's ability to discriminate leaping notes.

Indices of Difficulty and Discrimination. The difficulty indices for item 15 was .34. The point-biserial index of correlation, .54, was within suggested guidelines.

Items 17 and 18

Items 17 and 18 were designed to assess a student's ability to identify notes by their absolute letter names (A, B, C, D, E, F, and G).

Indices of Difficulty and Discrimination. The difficulty indices for items 17 and 18 were .61 and .68. Point-biserial indices of discrimination were .57 and .59.

Items 19 and 20

Items 19 and 20 were designed to assess a student's ability to equate notes with their corresponding solfa syllable names (do, re, mi, fa, so, la, and ti).

Indices of Difficulty and Discrimination. Item 19 ($p = .39$) was more difficult than item 20 ($p = .50$). Both items discriminated positively (item 19: $r_{pbis} = .68$; item 20: $r_{pbis} = .62$).

APPENDIX N

QUESTIONNAIRE COMPLETED BY MUSIC TEACHERS: ILLUSTRATION OF STRUCTURE

Students Opportunity to Learn
Test Version #6

The content of the music achievement test developed in this study is based upon an examination of selected textbook series and methodologies in which music literacy is viewed, at least in part, as the ability to read and/or write music notation. More specifically, this analysis involved the music education series Music and You (1988) and Musicanada (1983), the Kodaly method as interpreted in The Kodaly Method (Choksy, 1988) and the Orff process as interpreted in Exploring Orff (Steen, 1992). In addition, the music curriculum developed by the Province of Alberta (1989) was consulted. As a result of this survey, melodic perception was subdivided into ten characteristics of melody.

Due to limited class time, it is not always possible to progress through any particular music curriculum at the prescribed rate. Teachers are responsible for creating individual courses which best meet the needs of their students in the time available for instruction. Therefore, the emphasis placed on different characteristics of melodic perception varies across programs. As a result, no attempt was made in this study to assess the relative importance of the ten characteristics identified in the survey of the literature. For the most part, each is examined by means of two items.

Since this test deals specifically with music achievement, the amount of class time devoted to each characteristic of melody will most likely influence the success with which students complete the corresponding items. Therefore, I need your assistance in order to more accurately interpret the difficulty of individual items. The information you provide will be gathered through the accompanying questionnaire.

The questionnaire is composed of twelve independent segments. Each segment begins with a statement of the characteristic of melody which was identified in the survey described above. This is followed by an accompanying test item. You are asked three questions in each segment. The first relates to your students' opportunity to engage in musical activities associated with the characteristic of melody described within that segment. The second relates to your students' opportunity to engage in musical activities similar to the requirements of the accompanying test item. For segments in which students have had the opportunity to engage in musical activities related to the test item, you are asked to indicate how recently this characteristic of melody was discussed in class.

Instructions for Response

Each of the 12 questionnaire segments is divided into two parts.

Part 1- Characteristic of Melody

Sample Characteristic of Melody: Melodies have a tonal center.

Question 1:

In your estimation, what is the extent to which students in your music program have had the opportunity to engage in musical activities which relate to the characteristic of melody described above?

Please check (√) the appropriate box.

RATING	FULL OPPORTUNITY	SOME OPPORTUNITY	NO OPPORTUNITY
RESPONSE			

Guidelines for responses are described below:

FULL OPPORTUNITY: The students have been introduced to this characteristic of melody and have often had a full opportunity to engage in follow-up activities related to this characteristic of melodic perception.

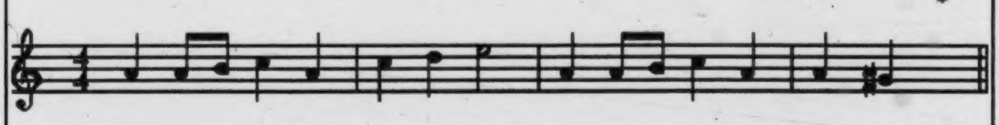
SOME OPPORTUNITY: The students have been introduced to this characteristic of melody and have had some opportunity to engage in follow-up activities related to this characteristic of melodic perception.

NO OPPORTUNITY: The students have had no opportunity to engage in activities related to this characteristic of melody.

Part 2
Test Items

Sample Test Item

Listen to the melody once. The last note is missing. Complete the melody so that it ends on the **home tone** or **tonal center**. Write your answer in the correct place on the staff below the star (*).



Question 2:

In your estimation, what is the extent to which students in your music program have had the opportunity to engage in musical activities similar to the requirements of the sample test item?
Please check (√) the appropriate box.

RATING	FULL OPPORTUNITY	SOME OPPORTUNITY	NO OPPORTUNITY
RESPONSE			

Guidelines for responses are described below:

FULL OPPORTUNITY: The students have been introduced to musical activities similar to the requirements of the test item and have had a full opportunity to engage in follow-up activities related to this characteristic of melodic perception.

SOME OPPORTUNITY: The students have been introduced to musical activities similar to the requirements of the test item and have had some opportunity to engage in follow-up activities.

NO OPPORTUNITY: The students have had no opportunity to engage in musical activities similar to the requirements of the test item.

IF STUDENTS HAVE HAD EITHER A 'FULL OPPORTUNITY' OR 'SOME OPPORTUNITY,' TO ENGAGE IN FOLLOW-UP ACTIVITIES, CONTINUE WITH QUESTION 3.

QUESTION 3:

When did students last have the opportunity to engage in follow-up activities similar to the requirements of the sample test item?

Guidelines for response:
This is an open-ended question. Examples of response might be- last music class, 2 weeks ago, 1 month ago, 2 years ago.

____

MATERIALS FOR TEACHING:

IN BRIEF, PLEASE DESCRIBE THE BOOKS AND MATERIALS YOU USE IN LESSON PREPARATION:

____

____

____

____

____

____

____

____

____

____

____

____

Note: Using the questionnaire structure illustrated here music teachers assessed items 1, 3, 5, 7, 9, 12, 13, 15, 18, 19, and 22.

APPENDIX O

FORMAL MUSIC EXPERIENCE

Field Study 6: Formal Music Experience: Summary of Gender and Instruments Played.

Student	Gender	Instrument
1	M	piano
2	F	piano
3	F	piano
4	F	piano, voice
5	F	piano
6	F	piano, clarinet, violin
7	F	piano, voice
8	M	piano
9	F	piano
10	M	piano
11	F	violin
12	F	piano
13	M	piano
14	M	cello
15	F	piano
16	F	piano
17	F	piano
18	M	piano
19	F	piano
20	M	piano
21	M	violin
22	F	piano
23	F	voice
24	M	piano
25	M	piano
26	F	piano
27	F	guitar
28	M	piano
29	F	piano
30	F	flute
31	F	piano

Field Study 6: Formal Music Experience: Summary of Gender and Instruments Played.

Student	Gender	Instrument
32	M	piano
33	M	piano
34	F	piano
35	M	piano
36	F	piano
37	F	piano
38	F	piano
39	F	saxophone
40	F	piano
41	F	piano
42	F	piano
43	F	piano
44	M	piano
45	F	piano
46	F	piano
47	M	guitar
48	M	piano
49	M	bagpipes
50	F	voice
51	M	piano
52	M	piano
53	F	piano
54	M	piano
55	M	piano
56	F	piano
57	F	piano
58	M	piano
59	F	piano
60	F	piano
61	F	piano
62	F	piano

Field Study 6: Formal Music Experience: Summary of Gender and Instruments Played.

Student	Gender	Instrument
63	M	piano
64	M	piano
65	F	piano, violin
66	M	piano
67	M	piano, violin
68	F	piano
69	M	violin
70	F	piano
71	F	piano
72	F	piano
73	M	piano
74	M	violin
75	F	piano
76	M	guitar
77	M	piano
78	F	piano
79	F	piano, violin
80	M	piano
81	M	piano, violin
82	M	piano, violin
83	F	piano, cello
84	F	piano
85	F	piano
86	F	piano
87	M	piano
88	F	piano
89	F	piano, violin, voice
90	F	piano
91	M	piano
92	M	piano
93	M	piano

Field Study 6: Formal Music Experience: Summary of Gender and Instruments Played.

Student	Gender	Instrument
94	M	piano
95	M	piano
96	F	piano
97	F	piano
98	M	piano
99	M	piano
100	M	piano

APPENDIX P

FACTOR STRUCTURE: PATTERN MATRIX

Factor Analysis: Rotated to Obliman Solution (Pattern Matrix) to Approximate Simple Structure.

	Factors		
	1	2	3
Items	Aural/Visual	Visual	Contour
1	.50917	-.09376	.04984
2	.47537	.03981	-.03355
3	.51044	-.29353	.08614
4	.63720	.23208	-.09253
5	.59689	-.14071	-.06831
6	.46990	-.26529	.13883
7	.48744	-.12471	-.02866
8	-.24892	.01956	.13751
9	.05673	-.11341	.67209
10	.10542	.07296	-.45852
11	.37624	-.08723	-.36162
12	-.41125	.25905	.21594
13	.42460	.31069	.24152
14	-.20014	-.42159	-.30998
15	.10924	-.58801	-.00095
16	-.13009	-.62876	.04025
17	.12287	-.68577	.22935
18	.17041	-.67065	.18116
19	.21705	-.71279	.01334
20	.32064	-.52069	-.02518
21	.00814	.17004	-.41562
22	.11864	-.23655	-.51436

APPENDIX Q

ETHICS APPROVAL

September 21, 1994

Ms Sheila J. Scott
405-10405 Saskatchewan Drive
EDMONTON, Alberta
T6E 4R9

Dear Ms. Scott:

Thank you for faxing your letter and ethics review approval. We would be pleased to have you complete this worthwhile project in music achievement in our jurisdiction. I will forward a copy of my letter of approval to the principals of the respective schools named and you may contact them directly to make further arrangements. I believe we would appreciate knowing the results of your work.

Sincerely,

Mr. Randy Wimmer, Director
Practicum Placement
Division of Field Services
University of Alberta
Edmonton, Alberta
T6G 2G5

Dear Mr. Wimmer:

Re: Research Request: Development of a Test Assessing Music Literacy Skills in the Upper Elementary Grades: (de Frece)

The above research request has been approved on a permissive basis following examination by our department. The approval is subject to the following conditions.

1. Teachers and student participation in the study to be voluntary;
2. Students are free to withdraw at any time;
3. Parental permission will be sought for students to participate in the study;
4. The results of the study will be provided to the teacher;
5. Anonymity of the students and the confidentiality of information obtained is assured; and
6. The researcher provides a copy of the results to this office.

Sheila Scott should now contact the undernoted principals to obtain approval and to make the necessary arrangements for conducting the study.

The district would appreciate receiving a copy of the study results as they become available. **(PLEASE QUOTE FILE # ABOVE)**

Yours sincerely,

I, verify that Sheila Scott had permission to conduct research in connection with the study "The construction and revision of test items referenced to selected fifth-grade music reading abilities" at Elementary School during the 1994-1995 school year.

Signature

Principal
Title

June 27, 1996
Date